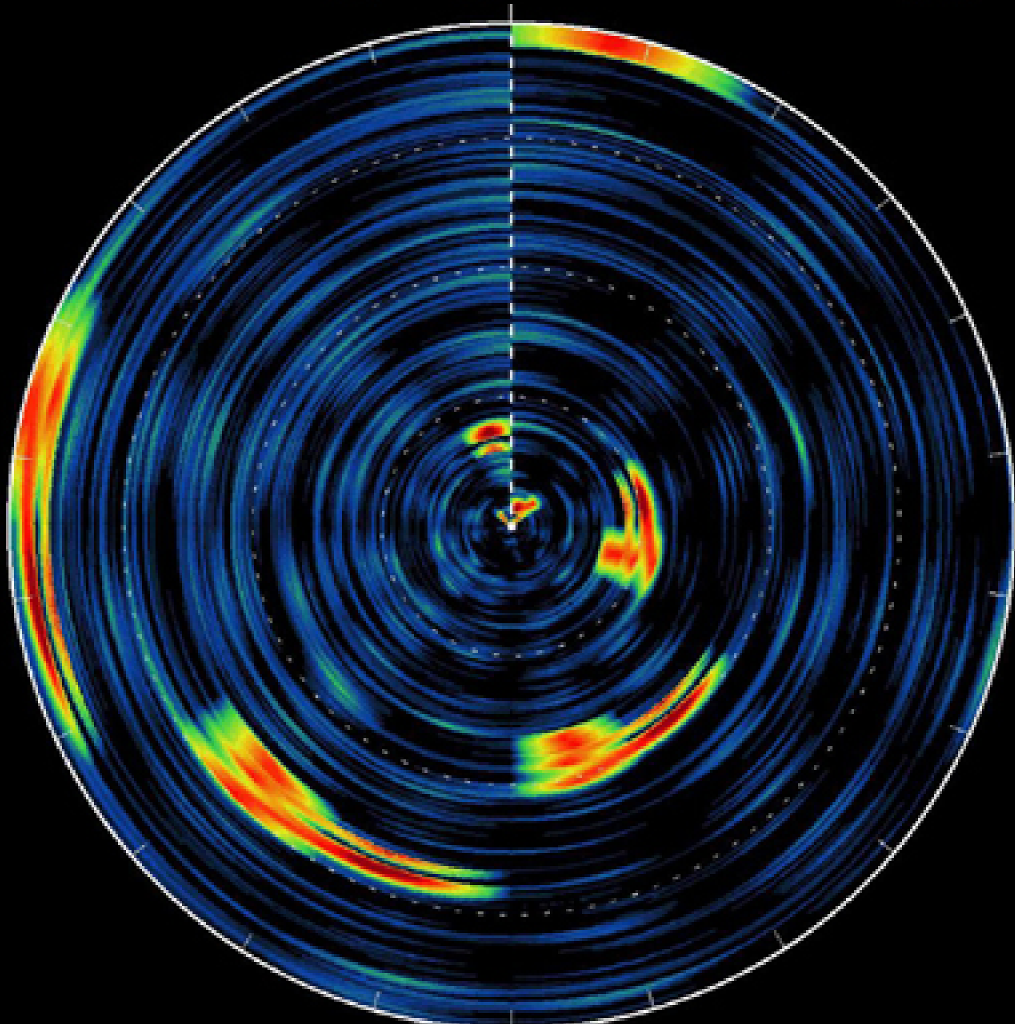
Radar Signal Processing, Sound Navigation and Ranging Technology



Jonelle Loper

Mallie Blocker

First Edition, 2012

ISBN 978-81-323-0975-8

Published by:
Academic Studio
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter-1

Digital Signal Processing

Digital signal processing (**DSP**) is concerned with the representation of signals by a sequence of numbers or symbols and the processing of these signals. Digital signal processing and analog signal processing are subfields of signal processing. DSP includes subfields like: audio and speech signal processing, sonar and radar signal processing, sensor array processing, spectral estimation, statistical signal processing, digital image processing, signal processing for communications, control of systems, biomedical signal processing, seismic data processing, etc.

The goal of DSP is usually to measure, filter and/or compress continuous real-world analog signals. The first step is usually to convert the signal from an analog to a digital form, by *sampling* it using an analog-to-digital converter (ADC), which turns the analog signal into a stream of numbers. However, often, the required output signal is another analog output signal, which requires a digital-to-analog converter (DAC). Even if this process is more complex than analog processing and has a discrete value range, the application of computational power to digital signal processing allows for many advantages over analog processing in many applications, such as error detection and correction in transmission as well as data compression.

DSP algorithms have long been run on standard computers, on specialized processors called digital signal processors (DSPs), or on purpose-built hardware such as application-specific integrated circuit (ASICs). Today there are additional technologies used for digital signal processing including more powerful general purpose microprocessors, field-programmable gate arrays (FPGAs), digital signal controllers (mostly for industrial apps such as motor control), and stream processors, among others.

Signal sampling

With the increasing use of computers the usage of and need for digital signal processing has increased. In order to use an analog signal on a computer it must be digitized with an analog-to-digital converter. Sampling is usually carried out in two stages, discretization and quantization. In the discretization stage, the space of signals is partitioned into equivalence classes and quantization is carried out by replacing the signal with representative signal of the corresponding equivalence class. In the quantization stage the representative signal values are approximated by values from a finite set.

The Nyquist–Shannon sampling theorem states that a signal can be exactly reconstructed from its samples if the sampling frequency is greater than twice the highest frequency of the signal; but requires an infinite number of samples . In practice, the sampling frequency is often significantly more than twice that required by the signal's limited bandwidth.

A digital-to-analog converter is used to convert the digital signal back to analog. The use of a digital computer is a key ingredient in digital control systems.

DSP domains

In DSP, engineers usually study digital signals in one of the following domains: time domain (one-dimensional signals), spatial domain (multidimensional signals), frequency domain, autocorrelation domain, and wavelet domains. They choose the domain in which to process a signal by making an informed guess (or by trying different possibilities) as to which domain best represents the essential characteristics of the signal. A sequence of samples from a measuring device produces a time or spatial domain representation, whereas a discrete Fourier transform produces the frequency domain information, that is the frequency spectrum. Autocorrelation is defined as the cross-correlation of the signal with itself over varying intervals of time or space.

Time and space domains

The most common processing approach in the time or space domain is enhancement of the input signal through a method called filtering. Digital filtering generally consists of some linear transformation of a number of surrounding samples around the current sample of the input or output signal. There are various ways to characterize filters; for example:

- A "linear" filter is a linear transformation of input samples; other filters are "non-linear". Linear filters satisfy the superposition condition, i.e. if an input is a weighted linear combination of different signals, the output is an equally weighted linear combination of the corresponding output signals.

- A "causal" filter uses only previous samples of the input or output signals; while a "non-causal" filter uses future input samples. A non-causal filter can usually be changed into a causal filter by adding a delay to it.

- A "time-invariant" filter has constant properties over time; other filters such as adaptive filters change in time.

- Some filters are "stable", others are "unstable". A stable filter produces an output that converges to a constant value with time, or remains bounded within a finite interval. An unstable filter can produce an output that grows without bounds, with bounded or even zero input.

- A "finite impulse response" (FIR) filter uses only the input signals, while an "infinite impulse response" filter (IIR) uses both the input signal and previous samples of the output signal. FIR filters are always stable, while IIR filters may be unstable.

Filters can be represented by block diagrams which can then be used to derive a sample processing algorithm to implement the filter using hardware instructions. A filter may also be described as a difference equation, a collection of zeroes and poles or, if it is an FIR filter, an impulse response or step response.

The output of a digital filter to any given input may be calculated by convolving the input signal with the impulse response.

Frequency domain

Signals are converted from time or space domain to the frequency domain usually through the Fourier transform. The Fourier transform converts the signal information to a magnitude and phase component of each frequency. Often the Fourier transform is converted to the power spectrum, which is the magnitude of each frequency component squared.

The most common purpose for analysis of signals in the frequency domain is analysis of signal properties. The engineer can study the spectrum to determine which frequencies are present in the input signal and which are missing.

In addition to frequency information, phase information is often needed. This can be obtained from the Fourier transform. With some applications, how the phase varies with frequency can be a significant consideration.

Filtering, particularly in non-realtime work can also be achieved by converting to the frequency domain, applying the filter and then converting back to the time domain. This is a fast, O(n log n) operation, and can give essentially any filter shape including excellent approximations to brickwall filters.

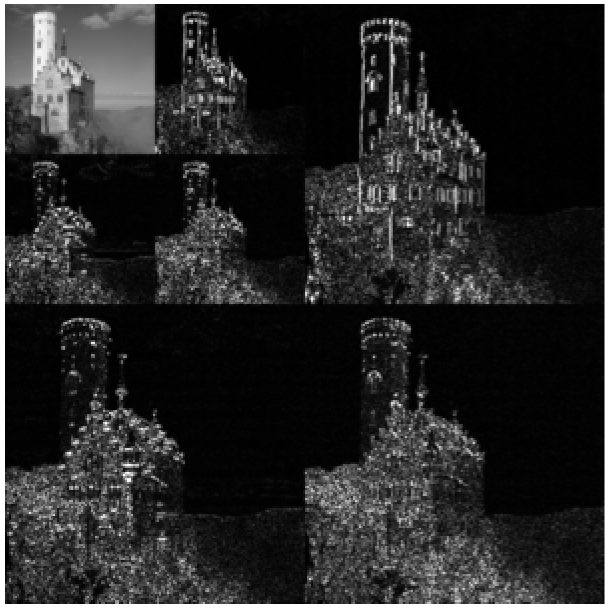
There are some commonly used frequency domain transformations. For example, the cepstrum converts a signal to the frequency domain through Fourier transform, takes the logarithm, then applies another Fourier transform. This emphasizes the frequency components with smaller magnitude while retaining the order of magnitudes of frequency components.

Frequency domain analysis is also called *spectrum-* or *spectral analysis*.

Z-plane analysis

Whereas analog filters are usually analysed in terms of transfer functions in the s plane using Laplace transforms, digital filters are analysed in the z plane in terms of Z-transforms. A digital filter may be described in the z plane by its characteristic collection of zeroes and poles.

Wavelet



An example of the 2D discrete wavelet transform that is used in JPEG2000. The original image is high-pass filtered, yielding the three large images, each describing local changes in brightness (details) in the original image. It is then low-pass filtered and downscaled, yielding an approximation image; this image is high-pass filtered to produce the three smaller detail images, and low-pass filtered to produce the final approximation image in the upper-left.

In numerical analysis and functional analysis, a **discrete wavelet transform** (DWT) is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency *and* location information (location in time).

Applications

The main applications of DSP are audio signal processing, audio compression, digital image processing, video compression, speech processing, speech recognition, digital communications, RADAR, SONAR, seismology and biomedicine. Specific examples are speech compression and transmission in digital mobile phones, room correction of sound in hi-fi and sound reinforcement applications, weather forecasting, economic forecasting, seismic data processing, analysis and control of industrial processes, medical imaging such as CAT scans and MRI, MP3 compression, computer graphics, image manipulation, hi-fi loudspeaker crossovers and equalization, and audio effects for use with electric guitar amplifiers.

Implementation

Digital signal processing is often implemented using specialised microprocessors such as the DSP56000, the TMS320, or the SHARC. These often process data using fixed-point arithmetic, although some versions are available which use floating point arithmetic and are more powerful. For faster applications FPGAs might be used. Beginning in 2007, multicore implementations of DSPs have started to emerge from companies including Freescale and Stream Processors, Inc. For faster applications with vast usage, ASICs might be designed specifically. For slow applications, a traditional slower processor such as a microcontroller may be adequate. Also a growing number of DSP applications are now being implemented on Embedded Systems using powerful PCs with a Multi-core processor.

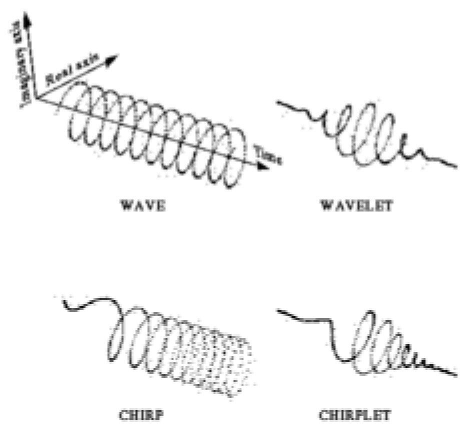
Techniques

- Bilinear transform
- Discrete Fourier transform
- Discrete-time Fourier transform
- Filter design
- LTI system theory
- Minimum phase
- Transfer function
- Z-transform
- Goertzel algorithm
- s-plane

Chapter-2

Chirplet Transform and Electronic Counter-countermeasures

Chirplet transform



Comparison of wave, wavelet, chirp, and chirplet

In signal processing, the **chirplet transform** is an inner product of an input signal with a family of analysis primitives called **chirplets**.

Similarity to other transforms

Much as in the wavelet transform, the chirplets are usually generated from (or can be expressed as being from) a single *mother chirplet* (analogous to the so-called *mother wavelet* of wavelet theory).

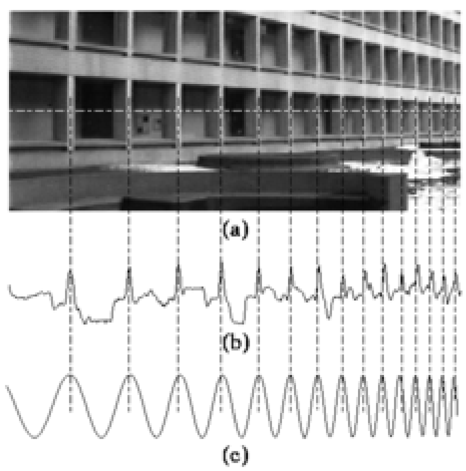
What is the chirplet and chirplet transform?

The term *chirplet transform* was coined by Steve Mann, as the title of the first published paper on chirplets. The term *chirplet* itself (apart from chirplet transform) was also used by Steve Mann, Domingo Mihovilovic, and Ronald Bracewell to describe a windowed portion of a chirp function. In Mann's words:

A wavelet is a piece of a wave, and a chirplet, similarly, is a piece of a chirp. More precisely, a chirplet is a windowed portion of a chirp function, where the window provides some time localization property. In terms of time–frequency space, chirplets exist as rotated, sheared, or other structures that move from the traditional parallelism with the time and frequency axes that are typical for waves (Fourier and short-time Fourier transforms) or wavelets.

The chirplet transform thus represents a rotated, sheared, or otherwise transformed tiling of the time–frequency plane. Although chirp signals have been known for many years in radar, pulse compression, and the like, the first published reference to the *chirplet transform* described specific signal representations based on families of functions related to one another by time–varying frequency modulation or frequency varying time modulation, in addition to time and frequency shifting, and scale changes. In that paper, the Gaussian chirplet transform was presented as one such example, together with a successful application to ice fragment detection in radar (improving target detection results over previous approaches). The term *chirplet* (but not the term *chirplet transform*) was also proposed for a similar transform, apparently independently, by Mihovilovic and Bracewell later that same year.

Applications



(a) In image processing, periodicity is often subject to a linear scaling. (b) In this image, repeating structures like the alternating dark space inside the windows, and light space of the white concrete, *chirp* (increase in frequency) towards the right. (c) The chirplet transform is able to represent this modulated variation compactly.

The chirplet transform is a useful signal analysis and representation framework that is widely used in

- Radar;
- Biomedical, the most commonly and widely used chirplet applications in biomedical being:
    - Heart sound processing;

- EEG processing.
- Signal processing;
- Image processing;
- SETI@home uses chirp functions to compensate for Doppler drift;
- Chirplet Time Domain Reflectometry.

Taxonomy of chirplet transforms

There are two broad categories of chirplet transform:

- Fixed
- Adaptive

These categories may be further subdivided by:

- choice of chirp
- choice of window

In either the fixed or adaptive case, the chirplets may be:

- q-chirplets (quadratic chirplets) of the form $\exp[i2\pi(at^2 + bt + c)]$ or, in general, some kind of quadratically varying exponent, linear swept wave packet, or the like. These are sometimes called *linear FM chirplets* (linear frequency-modulated chirplets, since quadratic phase is linear frequency). Commonly used families of q-chirplets are metaplectomorphisms of one another (i.e., the energy distribution of any member of the family of q-chirplets can be generated from any other member by shear-in-time, shear-in-frequency, dilation, translation-in-time, and translation-in-frequency).
- w-chirplets, also known as *warblets*. A family of warblets are like the sound made by birds called warblers. Unwindowed warblets have a sinusoidally varying time–frequency distribution, or similar cyclostationary or periodically varying time–frequency plot. The sound of a police siren is an example, in which the pitch goes up and down periodically. Of course, the warblet is a "piece of" a warble (i.e., a windowed section of something that has a time–frequency periodicity).
- d-chirplets, also known as *Doppler chirplets*. These are analysis functions that mimic the Doppler shift of a passing tone (e.g., the sound you hear from a train whistle as it moves past).
- p-chirplets, in which the scale varies projectively. Whereas the wavelet transform is based on wavelets of the form $g(ax + b)$, the p-type chirplet transform is based on chirplets of the form $g((ax + b)/(cx + 1))$, where a is the scale, b is the translation, and c is the *chirpiness* (chirp-rate, as defined by the degree of perspective, or projection).

The choice of window is also another matter of decision. A Gaussian window is one possible choice, leading to a four parameter chirplet transform (for which time–shear and frequency–shear only give one degree of freedom that may thus be encapsulated as rotation angle—Radon transform of the Wigner distribution may, for example, be used, as may the fractional Fourier transform).

Another possible choice is the rectangular window, and discrete prolate spheroidal sequences (also called Multitaper#The_Slepian_sequences) may be used, by way of the *method of multiple mother chirplets*. This method gives a total chirplet transform as the sum of energies in various contributory chirplet transforms made from multiple windows, akin to the way in which DPSSs are used to get a perfect rectangular tiling of the time–frequency plane. Thus it is now possible to get perfect parallelogram tiling of the time–frequency plane, using the method of multiple mother chirplets.

Related work

The chirplet transform is a generalized representation that includes as special cases:

- The Fourier transform
- The short-time Fourier transform (STFT), also known as the spectrogram
- The Wigner-Ville distribution
- The wavelet transform
- Canonical conjugate variables
- Segal–Shale–Weil distribution

Josef Segman proposed the idea of incorporating scale into the Heisenberg group (position, momentum, phase, or equivalently any canonical conjugate variables taken together with phase, such as, for example, time, frequency, and phase). This gave rise to a four parameter space of time, frequency, phase, and scale. Segman introduced this idea of *phase scale*. (Personal communication with Mann, from Josef Segman, at Harvard University and at Massachusetts Institute of Technology). Further personal communication between Irving Segal (the principal behind the Segal, Shale Weil representation, known also as the metaplectic representation—a double covering of the symplectic group) and Mann led to additional insight into the chirplet transform, in particular, to the variation of the chirplet transform that is based on q-chirplets.

Electronic counter-countermeasures

Electronic counter-countermeasures (ECCM) is a part of electronic warfare which includes a variety of practices which attempt to reduce or eliminate the effect of electronic countermeasures (ECM) on electronic sensors aboard vehicles, ships and aircraft and weapons such as missiles. ECCM is also known as **electronic protective measures** (EPM), chiefly in Europe. In practice, EPM often means resistance to jamming.

History

Ever since electronics have been used in battle in an attempt to gain superiority over the enemy, effort has been spent on techniques to reduce the effectiveness of those electronics. More recently, sensors and weapons are being modified to deal with this threat. One of the most common types of ECM is radar jamming or spoofing. This originated with the Royal Air Force use of what they codenamed *window* during World War II, which is now often referred to as *chaff*. Jamming also may have originated with the British during World War II, when they began jamming German radio communications.

In perhaps the first example of ECCM, the Germans increased their radio transmitter power in an attempt to 'burn through' or override the British jamming, which by necessity of the jammer being airborne or further away produced weaker signals. This is still one of the primary methods of ECCM today. For example, modern airborne jammers are able to identify incoming radar signals from other aircraft and send them back with random delays and other modifications in an attempt to confuse the opponent's radar set, making the 'blip' jump around wildly and be impossible to range. More powerful airborne radars means that it is possible to 'burn through' the jamming at much greater ranges by overpowering the jamming energy with the actual radar returns. The Germans were not really able to overcome the chaff spoofing very successfully and had to work around it (by guiding the aircraft to the target area and then having them visually acquire the targets).

Today, more powerful electronics with smarter software for operation of the radar might be able to better discriminate between a moving target like an aircraft and an almost stationary target like a chaff bundle.

With the technology going into modern sensors and seekers, it is inevitable that all successful systems have to have ECCM designed into them, lest they become useless on the battlefield. In fact, the 'electronic battlefield' is often used to refer to ECM, ECCM and ELINT activities, indicating that this has become a secondary battle in itself.

Specific ECCM techniques

The following are some examples of EPM (other than simply increasing the fidelity of sensors through techniques such as increasing power or improving discrimination):

ECM detection

Sensor logic may be programmed to be able to recognize attempts at spoofing (e.g., aircraft dropping chaff during terminal homing phase) and ignore them. Even more sophisticated applications of ECCM might be to recognize the type of ECM being used, and be able to cancel out the signal.

Pulse compression by "chirping", or linear frequency modulation

One of the effects of the pulse compression technique, is boosting the apparent signal strength as perceived by the radar receiver. The outgoing radar pulses are chirped, that is, the frequency of the carrier is varied within the pulse, much like the sound of a cricket chirping. When the pulse reflects off a target and returns to the receiver, the signal is processed to add a delay as a function of the frequency. This has the effect of 'stacking' the pulse so it seems stronger, but shorter in duration, to further processors. The effect can increase the received signal strength to above that of noise jamming. Similarly, jamming pulses (used in deception jamming) will not typically have the same chirp, so will not benefit from the increase in signal strength.

Frequency hopping

Frequency agility ('frequency hopping') may be used to rapidly switch the frequency of the transmitted energy, and receiving only that frequency during the receiving time window. This foils jammers which cannot detect this frequency switch quickly enough, and switch their own jamming frequency accordingly during the receiving time window.

This method is also useful against barrage jamming, in that it forces the jammer to spread its jamming power across multiple frequencies in the jammed system's frequency range, reducing its power in the actual frequency used by the radar at any one time. The use of similar spread-spectrum techniques allow signals to be spread over a wide enough spectrum to make jamming of such a wideband signal difficult.

Sidelobe blanking

Radar jamming can be effective from directions other than the direction the radar antenna is currently aimed. When jamming is strong enough, the radar receiver can detect it from a relatively low gain sidelobe. The radar, however, will process signals as if they were received in the main lobe. Therefore, jamming can be seen in directions other than where the jammer is located. To combat this, an omnidirectional antenna is used for a comparison signal. By comparing the signal strength as received by both the omnidirectional and the (directional) main antenna, signals can be identified that are not from the direction of interest. These signals are then ignored.

Polarization

Polarization can be used to filter out unwanted signals, such as jamming. If a jammer and receiver do not have the same polarity, the jamming signal will incur a loss that reduces its effectiveness. The four basic polarities are horizontal, vertical, right-hand circular, and left-hand circular. The signal loss inherent in a cross polarized (transmitter different from receiver) pair is 3 decibels for dissimilar types, and 17 dB for opposites.

Aside from power loss to the jammer, radar receivers can also benefit from using two or more antennas of differing polarity and comparing the signals received on each. This effect can effectively eliminate all jamming of the wrong polarity, although enough jamming may still obscure the actual signal.

Radiation homing

The other main aspect of ECCM, is to program sensors or seekers to detect attempts at ECM and possible even to take advantage of it. For example, some modern fire-and-forget missiles like the Vympel R-77 and the AMRAAM are able to home in directly on sources of radar jamming if the jamming is too powerful to allow them to find and track the target normally. This mode, called 'home-on-jam', actually makes the missile's job easier. Some missile seekers actually target the enemy's radiation sources, and are therefore called "anti-radiation missiles" (ARM). The jamming in this case effectively becomes a beacon announcing the presence and location of the transmitter. This makes the use of such ECM a difficult decision; it may serve to obscure an exact location from a non-ARM missile, but in doing so it must emit signals which can be exploited by an ARM type missile.

Chapter-3

Matched Filter

In telecommunications, a **matched filter** (originally known as a **North filter**) is obtained by correlating a known signal, or template, with an unknown signal to detect the presence of the template in the unknown signal. This is equivalent to convolving the unknown signal with a conjugated time-reversed version of the template. The matched filter is the optimal linear filter for maximizing the signal to noise ratio (SNR) in the presence of additive stochastic noise. Matched filters are commonly used in radar, in which a known signal is sent out, and the reflected signal is examined for common elements of the out-going signal. Pulse compression is an example of matched filtering. Two-dimensional matched filters are commonly used in image processing, e.g., to improve SNR for X-ray pictures.

Derivation of the matched filter

The following section derives the matched filter for a discrete-time system. The derivation for a continuous-time system is similar, with summations replaced with integrals.

The matched filter is the linear filter, *h*, that maximizes the output signal-to-noise ratio.

$$y[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k].$$

Though we most often express filters as the impulse response of convolution systems, as above, it is easiest to think of the matched filter in the context of the inner product, which we will see shortly.

We can derive the linear filter that maximizes output signal-to-noise ratio by invoking a geometric argument. The intuition behind the matched filter relies on correlating the received signal (a vector) with a filter (another vector) that is parallel with the signal, maximizing the inner product. This enhances the signal. When we consider the additive stochastic noise, we have the additional challenge of minimizing the output due to noise by choosing a filter that is orthogonal to the noise.

Let us formally define the problem. We seek a filter, h, such that we maximize the output signal-to-noise ratio, where the output is the inner product of the filter and the observed signal x.

Our observed signal consists of the desirable signal s and additive noise v:

$$x = s + v.$$

Let us define the covariance matrix of the noise, reminding ourselves that this matrix has Hermitian symmetry, a property that will become useful in the derivation:

$$R_v = E\{vv^H\}$$

where $.^H$ denotes Hermitian (conjugate) transpose, and E denotes expectation. Let us call our output, y, the inner product of our filter and the observed signal such that

$$y = \sum_{k=-\infty}^{\infty} h^*[k]x[k] = h^H x = h^H s + h^H v = y_s + y_v.$$

We now define the signal-to-noise ratio, which is our objective function, to be the ratio of the power of the output due to the desired signal to the power of the output due to the noise:

$$SNR = \frac{|y_s|^2}{E\{|y_v|^2\}}.$$

We rewrite the above:

$$SNR = \frac{|h^H s|^2}{E\{|h^H v|^2\}}.$$

We wish to maximize this quantity by choosing h. Expanding the denominator of our objective function, we have

$$E\{|h^H v|^2\} = E\{(h^H v){(h^H v)}^H\} = h^H E\{vv^H\}h = h^H R_v h.$$

Now, our SNR becomes

$$SNR = \frac{|h^H s|^2}{h^H R_v h}.$$

We will rewrite this expression with some matrix manipulation. The reason for this seemingly counterproductive measure will become evident shortly. Exploiting the Hermitian symmetry of the covariance matrix R_v, we can write

$$SNR = \frac{|(R_v^{1/2}h)^H(R_v^{-1/2}s)|^2}{(R_v^{1/2}h)^H(R_v^{1/2}h)},$$

We would like to find an upper bound on this expression. To do so, we first recognize a form of the Cauchy-Schwarz inequality:

$$|a^Hb|^2 \leq (a^Ha)(b^Hb),$$

which is to say that the square of the inner product of two vectors can only be as large as the product of the individual inner products of the vectors. This concept returns to the intuition behind the matched filter: this upper bound is achieved when the two vectors *a* and *b* are parallel. We resume our derivation by expressing the upper bound on our *SNR* in light of the geometric inequality above:

$$SNR = \frac{|(R_v^{1/2}h)^H(R_v^{-1/2}s)|^2}{(R_v^{1/2}h)^H(R_v^{1/2}h)} \leq \frac{\left[(R_v^{1/2}h)^H(R_v^{1/2}h)\right]\left[(R_v^{-1/2}s)^H(R_v^{-1/2}s)\right]}{(R_v^{1/2}h)^H(R_v^{1/2}h)}.$$

Our valiant matrix manipulation has now paid off. We see that the expression for our upper bound can be greatly simplified:

$$SNR = \frac{|(R_v^{1/2}h)^H(R_v^{-1/2}s)|^2}{(R_v^{1/2}h)^H(R_v^{1/2}h)} \leq s^HR_v^{-1}s.$$

We can achieve this upper bound if we choose,

$$R_v^{1/2}h = \alpha R_v^{-1/2}s$$

where α is an arbitrary real number. To verify this, we plug into our expression for the output *SNR*:

$$SNR = \frac{|(R_v^{1/2}h)^H(R_v^{-1/2}s)|^2}{(R_v^{1/2}h)^H(R_v^{1/2}h)} = \frac{\alpha^2|(R_v^{-1/2}s)^H(R_v^{-1/2}s)|^2}{\alpha^2(R_v^{-1/2}s)^H(R_v^{-1/2}s)} = \frac{|s^HR_v^{-1}s|^2}{s^HR_v^{-1}s} = s^HR_v^{-1}s.$$

Thus, our optimal matched filter is

$$h = \alpha R_v^{-1}s.$$

We often choose to normalize the expected value of the power of the filter output due to the noise to unity. That is, we constrain

$$E\{|y_v|^2\} = 1.$$

This constraint implies a value of α, for which we can solve:

$$E\{|y_v|^2\} = \alpha^2 s^H R_v^{-1} s = 1,$$

yielding

$$\alpha = \frac{1}{\sqrt{s^H R_v^{-1} s}},$$

giving us our normalized filter,

$$h = \frac{1}{\sqrt{s^H R_v^{-1} s}} R_v^{-1} s.$$

If we care to write the impulse response of the filter for the convolution system, it is simply the complex conjugate time reversal of *h*.

Though we have derived the matched filter in discrete time, we can extend the concept to continuous-time systems if we replace R_v with the continuous-time autocorrelation function of the noise, assuming a continuous signal $s(t)$, continuous noise $v(t)$, and a continuous filter $h(t)$.

Alternative derivation of the matched filter

Alternatively, we may solve for the matched filter by solving our maximization problem with a Lagrangian. Again, the matched filter endeavors to maximize the output signal-to-noise ratio (*SNR*) of a filtered deterministic signal in stochastic additive noise. The observed sequence, again, is

$$x = s + v,$$

with the noise covariance matrix,

$$R_v = E\{vv^H\}.$$

The signal-to-noise ratio is

$$SNR = \frac{|y_s|^2}{E\{|y_v|^2\}}.$$

Evaluating the expression in the numerator, we have

$$|y_s|^2 = {y_s}^H y_s = h^H s s^H h.$$

and in the denominator,

$$E\{|y_v|^2\} = E\{{y_v}^H y_v\} = E\{h^H v v^H h\} = h^H R_v h.$$

The signal-to-noise ratio becomes

$$SNR = \frac{h^H s s^H h}{h^H R_v h}.$$

If we now constrain the denominator to be 1, the problem of maximizing *SNR* is reduced to maximizing the numerator. We can then formulate the problem using a Lagrange multiplier:

$$\begin{aligned}
&h^H R_v h = 1 \\
&\mathcal{L} = h^H s s^H h + \lambda(1 - h^H R_v h) \\
&\nabla_{h^*}\mathcal{L} = s s^H h - \lambda R_v h = 0 \\
&(s s^H) h = \lambda R_v h
\end{aligned}$$

which we recognize as an eigenvalue problem

$$h^H (s s^H) h = \lambda h^H R_v h = \lambda.$$

Since ss^H is of unit rank, it has only one nonzero eigenvalue. It can be shown that this eigenvalue equals

$$\lambda_{\max} = s^H R_v^{-1} s,$$

yielding the following optimal matched filter

$$h = \frac{1}{\sqrt{s^H R_v^{-1} s}} R_v^{-1} s.$$

This is the same result found in the previous section.

Frequency-domain interpretation

When viewed in the frequency domain, it is evident that the matched filter applies the greatest weighting to spectral components that have the greatest signal-to-noise ratio. Although in general this requires a non-flat frequency response, the associated distortion is not significant in situations such as radar and digital communications, where the original waveform is known and the objective is to detect the presence of this signal against the background noise.

Example of matched filter in radar and sonar

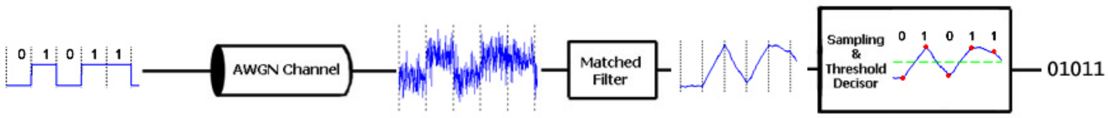
Matched filters are often used in signal detection. As an example, suppose that we wish to judge the distance of an object by reflecting a signal off it. We may choose to transmit a pure-tone sinusoid at 1 Hz. We assume that our received signal is an attenuated and phase-shifted form of the transmitted signal with added noise.

To judge the distance of the object, we correlate the received signal with a matched filter, which, in the case of white (uncorrelated) noise, is another pure-tone 1-Hz sinusoid. When the output of the matched filter system exceeds a certain threshold, we conclude with high probability that the received signal has been reflected off the object. Using the speed of propagation and the time that we first observe the reflected signal, we can estimate the distance of the object. If we change the shape of the pulse in a specially-designed way, the signal-to-noise ratio and the distance resolution can be even improved after matched filtering: this is a technique known as pulse compression.

Additionally, matched filters can be used in parameter estimation problems. To return to our previous example, we may desire to estimate the speed of the object, in addition to its position. To exploit the Doppler effect, we would like to estimate the frequency of the received signal. To do so, we may correlate the received signal with several matched filters of sinusoids at varying frequencies. The matched filter with the highest output will reveal, with high probability, the frequency of the reflected signal and help us determine the speed of the object. This method is, in fact, a simple version of the discrete Fourier transform (DFT). The DFT takes an N-valued complex input and correlates it with N matched filters, corresponding to complex exponentials at N different frequencies, to yield N complex-valued numbers corresponding to the relative amplitudes and phases of the sinusoidal components.

Example of matched filter in digital communications

The matched filter is also used in communications. In the context of a communication system that sends binary messages from the transmitter to the receiver across a noisy channel, a matched filter can be used to detect the transmitted pulses in the noisy received signal.



Imagine we want to send the sequence "0101100100" coded in non polar Non-return-to-zero (NRZ) through a certain channel.

Mathematically, a sequence in NRZ code can be described as a sequence of unit pulses or shifted rect functions, each pulse being weighted by +1 if the bit is "1" and by 0 if the bit is "0". Formally, the scaling factor for the k^{th} bit is,

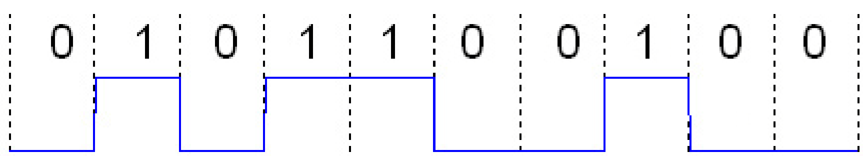
$$a_k = \begin{cases} 1, & \text{if bit } k \text{ is } 1, \\ 0, & \text{if bit } k \text{ is } 0. \end{cases}$$

We can represent our message, $M(t)$, as the sum of shifted unit pulses:

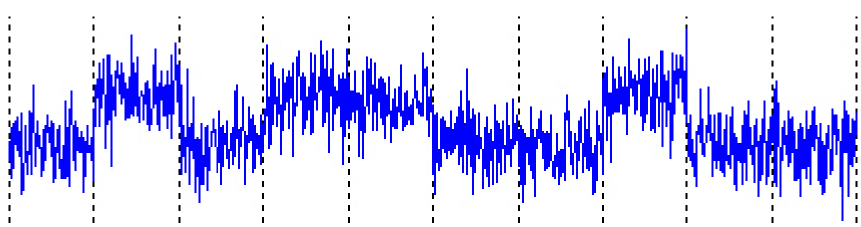
$$M(t) = \sum_{k=-\infty}^{\infty} a_k \times \Pi\left(\frac{t - kT}{T}\right).$$

where T is the time length of one bit.

Thus, the signal to be sent by the transmitter is



If we model our noisy channel as an AWGN channel, white Gaussian noise is added to the signal. At the receiver end, for a Signal-to-noise ratio of 3dB, this may look like:



A first glance will not reveal the original transmitted sequence. There is a high power of noise relative to the power of the desired signal (i.e., there is a low signal-to-noise ratio). If the receiver were to sample this signal at the correct moments, the resulting binary message would possibly belie the original transmitted one.

To increase our signal-to-noise ratio, we pass the received signal through a matched filter. In this case, the filter should be matched to an NRZ pulse (equivalent to a "1" coded in NRZ code). Precisely, the impulse response of the ideal matched filter, assuming white (uncorrelated) noise should be a time-reversed complex-conjugated scaled version of the signal that we are seeking. We choose

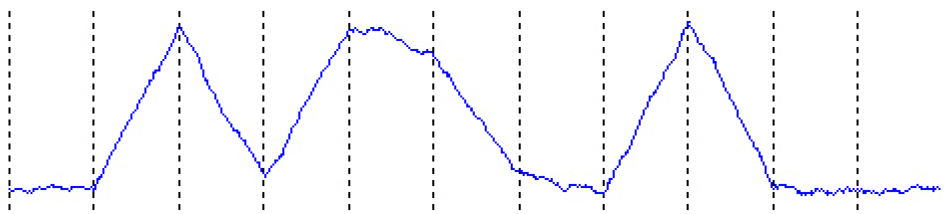
$$h(t) = \Pi\left(\frac{t}{T}\right).$$

In this case, due to symmetry, the time-reversed complex conjugate of $h(t)$ is in fact $h(t)$, allowing us to call $h(t)$ the impulse response of our matched filter convolution system.

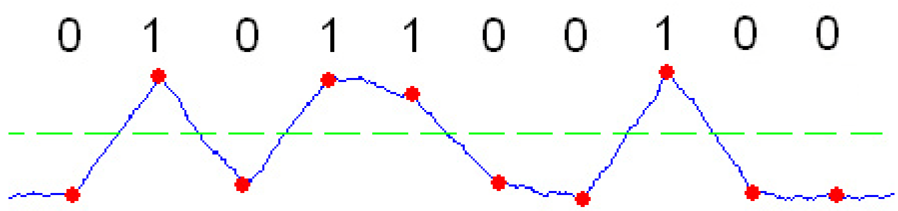
After convolving with the correct matched filter, the resulting signal, $M_{\text{filtered}}(t)$ is,

$$M_{\text{filtered}}(t) = M(t) * h(t)$$

where * denotes convolution.



Which can now be safely sampled by the receiver at the correct sampling instants, and compared to an appropriate threshold, resulting in a correct interpretation of the binary message.



Chapter-4

Pulse Compression

Pulse compression is a signal processing technique mainly used in radar, sonar and echography to augment the range resolution as well as the signal to noise ratio. This is achieved by modulating the transmitted pulse and then correlating the received signal with the transmitted pulse.

Simple pulse

Signal description

The simplest signal a pulse radar can transmit is a sinusoidal pulse of amplitude, $A$and carrier frequency, f_0, truncated by a rectangular function of width, T. The pulse is transmitted periodically, but that is not the main topic here; we will consider only a single pulse, s. If we assume the pulse to start at time $t=0$, the signal can be written the following way, using the complex notation:

$$s(t) = \begin{cases} Ae^{2i\pi f_0 t} & \text{if } 0 \leq t < T \\ 0 & \text{otherwise} \end{cases}$$

Range resolution

Let us determine the range resolution which can be obtained with such a signal. The return signal, written $r(t)$, is an attenuated and time-shifted copy of the original transmitted signal (in reality, Doppler effect can play a role too, but this is not important here.) There is also noise in the incoming signal, both on the imaginary and the real channel, which we will assume to be white and Gaussian (this generally holds in reality); we write $B(t)$to denote that noise. To detect the incoming signal, matched filtering is commonly used. This method is optimal when a known signal is to be detected among an additive white Gaussian noise.

In other words, the cross-correlation of the received signal with the transmitted signal is computed. This is achieved by convolving the incoming signal with a conjugated and time-reversed version of the transmitted signal. This operation can be done either in software or with hardware. We write $<s,r>(t)$for this cross-correlation. We have:

$$< s,r > (t) = \int_{t'=0}^{+\infty} s^{\star}(t')r(t+t')dt'$$

If the reflected signal comes back to the receiver at time $t_r$and is attenuated by factor K, this yields:

$$r(t) = \begin{cases} KAe^{2i\pi f_0(t-t_r)} + B(t) & \text{if } t_r \leq t < t_r + T \\ B(t) & \text{otherwise} \end{cases}$$

Since we know the transmitted signal, we obtain:

$$< s,r > (t) = KA^2\Lambda\left(\frac{t-t_r}{T}\right)e^{2i\pi f_0(t-t_r)} + B'(t)$$

where $B'(t)$, the result of the intercorrelation between the noise and the transmitted signal, remains a white noise of same characteristics as $B(t)$since it is not correlated to the transmitted signal. Function Λ is the triangle function, its value is 0 on $[-\infty, -\frac{1}{2}] \cup [\frac{1}{2}, +\infty]$, it augments linearly on $[-\frac{1}{2}, 0]$where it reaches its maximum 1, and it decreases linearly on $[0, \frac{1}{2}]$until it reaches 0 again. Figures at the end of this paragraph show the shape of the intercorrelation for a sample signal (in red), in this case a real truncated sine, of duration $T = 1$seconds, of unit amplitude, and frequency $f_0 = 10$hertz. Two echoes (in blue) come back with a delay of 3 and 5 seconds, respectively, and have an amplitude equal to 0.5 and 0.3; those are just random values for the sake of the example. Since the signal is real, the intercorrelation is weighted by an additional ½ factor.

If two pulses come back (nearly) at the same time, the intercorrelation is equal to the sum of the intercorrelations of the two elementary signals. To distinguish one "triangular" envelope from that of the other pulse, it is clearly visible that the times of arrival of the two pulses must be separated by at least Tso that the maxima of both pulses can be separated. If this condition is false, both triangles will be mixed together and impossible to separate.

Since the distance travelled by a wave during $T$is cT(where c is the speed of the wave in the medium), and since this distance corresponds to a round-trip time, we get:

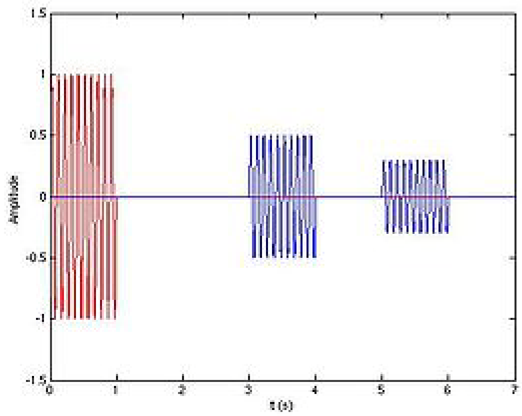
Result 1

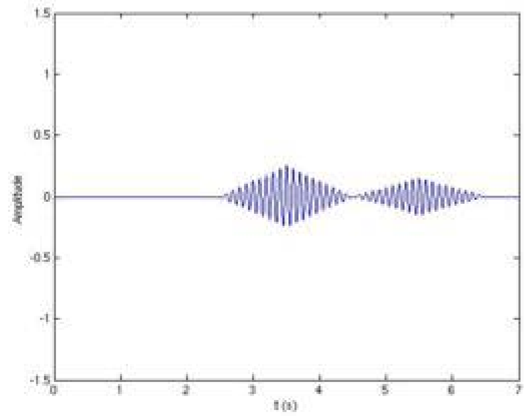
The range resolution with a sinusoidal pulse is $\frac{1}{2}cT$where $T$is the pulse Duration and, c, the speed of the wave.

Conclusion: to augment the resolution, the pulse length must be reduced.

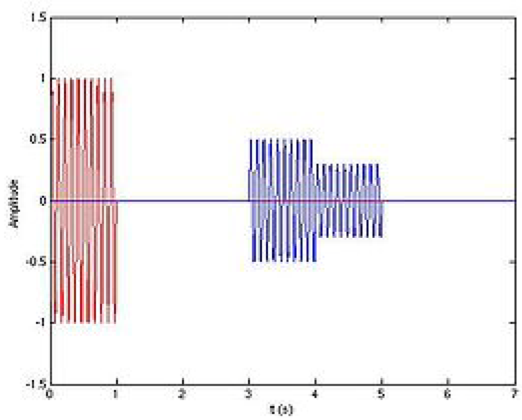
Example (simple impulsion): transmitted signal in red (carrier 10 hertz, amplitude 1, duration 1 second) and two echoes (in blue).

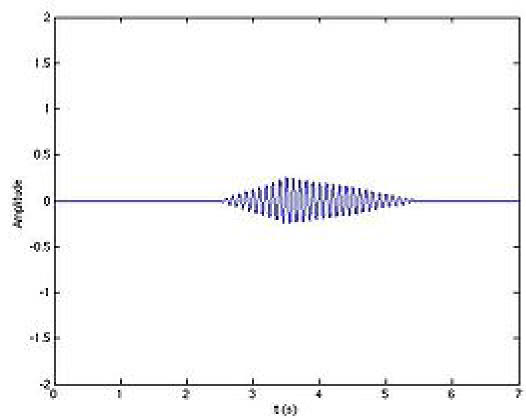
Before matched filtering | **After matched filtering**





If the targets are separated enough... ...echoes can be distinguished.





If the targets are too close... ...the echoes are mixed together.

Required energy to transmit that signal

The instantaneous power of the transmitted pulse is $P(t) = |s|^2(t)$. The energy put into that signal is:

$$E = \int_0^T P(t)dt = A^2T$$

Similarly, the energy in the received pulse is $E_r = K^2A^2T$. If $\sigma$is the standard deviation of the noise, the signal-to-noise ratio (SNR) at the receiver is:

$$SNR = \frac{E_r}{\sigma} = \frac{K^2 A^2 T}{\sigma}$$

The SNR augments with the pulse duration, if other parameters are frozen. This goes against the resolution requirements, since generally one wants a large resolution.

Pulse compression by linear frequency modulation (or chirping*)*

Basic principles

How can one have a large enough pulse (to still have a good SNR at the receiver) without poor resolution? This is where pulse compression enters the picture. The basic principle is the following:

- a signal is transmitted, with a long enough length so that the energy budget is correct
- this signal is designed so that after matched filtering, the width of the intercorrelated signals is smaller than the width obtained by the standard sinusoidal pulse, as explained above (hence the name of the technique: pulse compression).

In radar or sonar applications, linear chirps are the most typically used signals to achieve pulse compression. The pulse being of finite length, the amplitude is a rectangle function. If the transmitted signal has a duration T, begins at $t=0$and linearly sweeps the frequency band Δfcentered on carrier f_0, it can be written:

$$s_c(t) = \begin{cases} Ae^{2i\pi\left(f_0 + \frac{\Delta f}{2T}t - \frac{\Delta f}{2}\right)t} & \text{if } 0 \leq t < T \\ 0 & \text{otherwise} \end{cases}$$

N.b., the chirp is written that way so the phase of the chirped signal (that is, the argument of the complex exponential), is:

$$\phi(t) = 2\pi\left(f_0 + \frac{\Delta f}{2T}t - \frac{\Delta f}{2}\right)t$$

thus the instantaneous frequency is (by definition):

$$f(t) = \frac{1}{2\pi}\left[\frac{d\phi}{dt}\right]_t = f_0 - \frac{\Delta f}{2} + \frac{\Delta f}{T}t$$

which is the intended linear ramp going from $f_0 - \frac{\Delta f}{2}$at $t=0$to $f_0 + \frac{\Delta f}{2}$at $t=T$.

Cross-correlation between the transmitted and the received signal

As for the "simple" pulse, let us compute the cross-correlation between the transmitted and the received signal. To simplify things, we shall consider that the chirp is not written as it is given above, but in this alternate form (the final result will be the same):

$$s_{c'}(t) = \begin{cases} Ae^{2i\pi\left(f_0 + \frac{\Delta f}{2T}t\right)t} & \text{if } -\frac{T}{2} \le t < \frac{T}{2} \\ 0 & \text{else} \end{cases}$$

Since this cross-correlation is equal (save for the *K* attenuation factor), to the autocorrelation function of $s_{c'}$, this is what we consider:

$$< s_{c'}, s_{c'} > (t) = \int_{-\infty}^{+\infty} s_{c'}^{\star}(-t') s_{c'}(t - t') dt'$$

It can be shown that the autocorrelation function of $s_{c'}$ is:

$$< s_{c'}, s_{c'} > (t) = T\Lambda\left(\frac{t}{T}\right) sinc\left[\pi \Delta f t \Lambda\left(\frac{t}{T}\right)\right] e^{2i\pi f_0 t}$$

The maximum of the autocorrelation function of $s_{c'}$ is reached at 0. Around 0, this function behaves as the sinc term. The −3 dB temporal width of that cardinal sine is more or less equal to $T' = \frac{1}{\Delta f}$. Everything happens as if, after matched filtering, we had the resolution that would have been reached with a simple pulse of duration T'. For the common values of Δf, T' is smaller than T, hence the *pulse compression* name.

Since the cardinal sine can have annoying sidelobes, a common practice is to filter the result by a window (Hamming, Hann, etc.). In practice, this can be done at the same time as the adapted filtering by multiplying the reference chirp with the filter. The result will be a signal with a slightly lower maximum amplitude, but the sidelobes will be filtered out, which is more important.

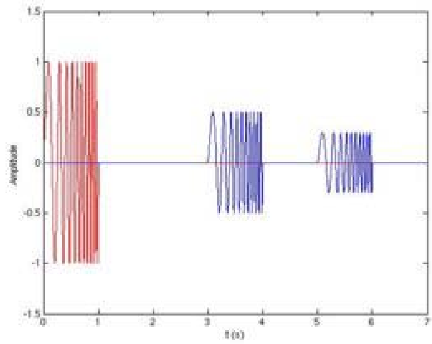
Result 2

The distance resolution reachable with a linear frequency modulation of a pulse on a bandwidth Δf is: $\frac{c}{2\Delta f}$ where c is the speed of the wave.

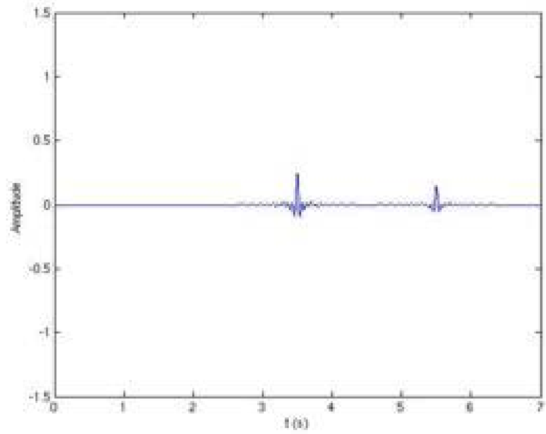
Definition

Ratio $\frac{T}{T'} = T\Delta f$ is the pulse compression ratio. It is generally greater than 1 (usually, its value is 20 to 30).

Example (chirped pulse): transmitted signal in red (carrier 10 hertz, modulation on 16 hertz, amplitude 1, duration 1 second) and two echoes (in blue).



Before matched filtering



After matched filtering: the echoes are shorter in time.

SNR augmentation through pulse compression

The energy of the signal does not vary during pulse compression. However, it is now located in the main lobe of the cardinal sine, whose width is approximately $T' \approx \frac{1}{\Delta f}$. If $P$is the power of the signal before compression, and $P'$the power of the signal after compression, we have:

$$P \times T = P' \times T'$$

which yields:

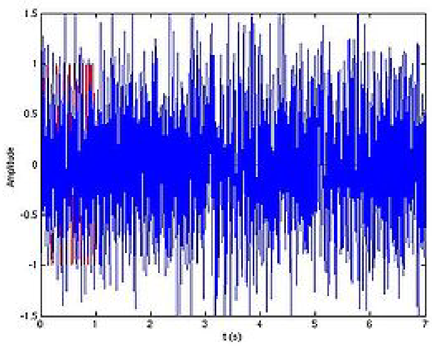
$$P' = P \times \frac{T}{T'}$$

Besides, the power of the noise does not change through intercorrelation since it is not correlated to the transmitted pulse (it is totally random). As a consequence:

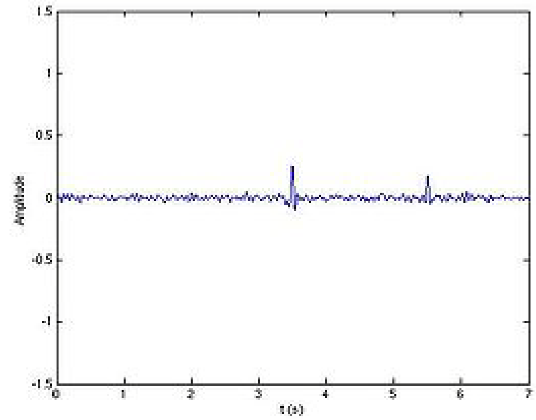
Result 3

After pulse compression, the power of the received signal can be considered as being amplified by $T\Delta f$. This additional gain can be injected in the radar equation.

Example: same signals as above, plus an additive Gaussian white noise ($\sigma = 0.5$)



Before matched filtering: the signal is hidden in noise



After matched filtering: echoes become visible.

Pulse compression by phase coding

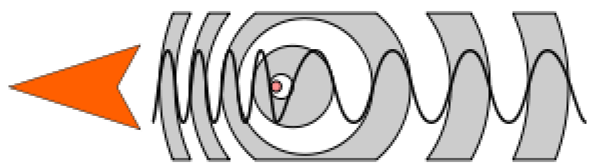
There are other means to modulate the signal. Phase modulation is a commonly used technique; in this case, the pulse is divided in N time slots of duration $\frac{T}{N}$ for which the phase at the origin is chosen according to a pre-established convention. For instance, it is possible not to change the phase for some time slots (which comes down to just leave the signal as it is, in those slots) and de-phase the signal in the other slots by π (which is equivalent of changing the sign of the signal). The precise way of choosing the sequence of $\{0, \pi\}$ phases is done according to a technique known as Barker codes. It is possible to code the sequence on more than two phases (polyphase coding). As with a linear chirp, pulse compression is achieved through intercorrelation.

The advantages of the Barker codes are their simplicity (as indicated above, a π de-phasing is a simple sign change), but the pulse compression ratio is lower than in the chirp case and the compression is very sensitive to frequency changes due to the Doppler effect if that change is larger than $\frac{1}{T}$.

Chapter-5

Doppler Radar and Moving Target Indication

Doppler radar



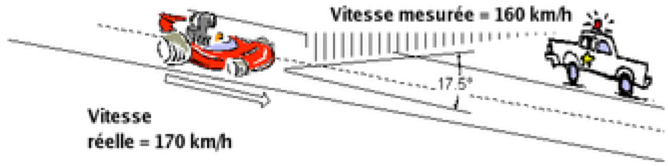
Doppler effect

A **Doppler radar** is a specialized radar that makes use of the doppler effect to produce velocity data about objects at a distance. It does this by beaming a microwave signal towards a desired target and listening for its reflection, then analyzing how the frequency of the returned signal has been altered by the object's motion. This variation gives direct and highly accurate measurements of the radial component of a target's velocity relative to the radar. Doppler radars are used in aviation, sounding satellites, meteorology, police speed guns, and radiology.

The specific term "*Doppler Radar*", due in part to its extremely common use by television meteorologists in on-air weather reporting, has erroneously become popularly synonymous with the type of radar used in meteorology. Most modern weather radars use the pulse-doppler technique to examine the motion of precipitation, but it is only a part of the processing of their data. So, while these radars use a highly specialized form of *doppler radar*, the term is much broader in its meaning and its applications.

Concept

Doppler Effect



The emitted signal toward the car is reflected back with a variation of frequency that depend on the speed away/toward the radar (160 km/h). This is only a component of the real speed (170 km/h).

The Doppler effect (or Doppler shift), named after Austrian physicist Christian Doppler who proposed it in 1842, is the change in frequency of a wave for an observer moving relative to the source of the waves. It is commonly heard when a vehicle sounding a siren approaches, passes and recedes from an observer. The received frequency is increased (compared to the emitted frequency) during the approach, it is identical at the instant of passing by, and it is decreased during the recession. This variation of frequency also depends on the direction the wave source is moving with respect to the observer; it is maximum when the source is moving directly toward or away from the observer, and diminishes with increasing angle between the direction of motion and the direction of the waves, until when the source is moving at right angles to the observer, there is no shift.

An analogy would be pitcher throwing one ball every second in a person's direction (a frequency of 1 ball per second). Assuming that the balls travel at a constant velocity, if the pitcher is stationary, the man will catch one ball every second. However, if the pitcher is jogging towards the man, he will catch balls more frequently because the balls will be less spaced out (the frequency increases). The inverse is true if the pitcher is moving away from the man; he will catch balls less frequently due to the pitcher's backward motion (the frequency decreases). If the pitcher were to move at an angle but with the same speed, the variation of the frequency at which the receiver would catch the ball would be less as the distance between the two would change more slowly.

Note that, from the point of view of the pitcher, the frequency remains constant (whether he's throwing balls or transmitting microwaves). Since with electromagnetic radiation like microwaves frequency is inversely proportional to wavelength, the wavelength of the waves is also affected. Thus, the relative difference in velocity between a source and an observer is what gives rise to the doppler effect.

Frequency variation

The exact formula for radar doppler shift is the same as that for reflection of light by a moving mirror. There is no need to invoke Einsteins' theory of special relativity, because all observations are made in the same frame of reference. The exact result derived with c as the speed of light and v as the target velocity gives the shifted frequency (F_r) as a function of the original frequency (F_t) :

$$F_r = F_t(\frac{1 + v/c}{1 - v/c})$$

The exact "beat frequency", aka Doppler Frequency (F_d), is thus:

$$F_d = F_r - F_t = 2v\frac{F_t}{(c - v)}$$

Since for most all practical applications of radar, $v \ll c$, so $(c - v) \rightarrow c$. We can then write:

$$F_d \approx 2v\frac{F_t}{c}$$

Technology

U.S. Army soldier using a radar gun, an application of Doppler radar, to catch speeding violators.

There are three ways of producing the Doppler effect. Radars may be Coherent pulsed (CP), Continuous wave (CW), or Frequency modulated (FM). CW doppler radar only provides a velocity output as the received signal from the target is compared in frequency with the original signal. Early doppler radars were CW, but these quickly led to the development of frequency modulated continuous wave(FM-CW) radar, which sweeps the transmitter frequency to encode and determine range.

The CW and FM-CW radars can normally only process one target, which limits their use. With the advent of digital techniques, Pulse-Doppler radars (PD) were introduced, and doppler processors for coherent pulse radars were developed at the same time. The advantage of combining doppler processing with pulse radars is to provide accurate velocity information. This velocity is called Range-Rate. It describes the rate that a target moves towards or away from the radar. A target with no range-rate reflects a frequency near the transmitter frequency, and cannot be detected. The classic zero doppler target is one which is on a heading that is tangential to the radar antenna beam. Basically, any target that is heading 90 degrees in relation to the antenna beam cannot be detected by its velocity (only by its conventional reflectivity).

History

FM radar was highly developed during World War II for the use by US Navy aircraft. Most used the UHF spectrum, and had a transmit yagi antenna on the port wing, and a receiver yagi antenna on the starboard wing. This allowed bombers to fly an optimum speed when approaching ship targets. Later when magnetrons and microwaves became available, the use of FM radar fell into disuse.

When the digital Fast Fourier transform became available, it was immediately connected to Coherent Pulsed radars, where velocity information was extracted. This quickly proved useful in both weather and air traffic control radars. The velocity information provided another input to the software tracker, and improved computer tracking. Due to the low *pulse repetition frequency* (PRF) of most coherent pulsed radars, which maximizes the coverage in range, the amount of doppler processing is limited. The doppler processor can only process velocities up to ±1/2 the PRF of the radar. This was not a problem for weather radars.

Specialized radars quickly were mechanized when digital techniques became affordable. Pulse-Doppler radars combine all the benefits of long range, and high velocity capability. Pulse-Doppler radars use a medium to high PRF (on the order of 30 kHz). This high PRF allows for the detection of either high speed targets, or high resolution velocity measurements. Normally it is one or the other, that is, a radar designed for detecting targets from zero to Mach 2, does not have a high resolution in speed, while a radar designed for high resolution velocity measurements does not have a wide range of speeds. Weather radars are high resolution velocity radars, while air defense radars have a large range of velocity detection, but the accuracy in velocity is in the 10's of knots.

Antenna designs for the CW and FM-CW started out as separate transmit and receive antennas before the advent of affordable microwave designs. In the late 1960s traffic radars began being produced which used a single antenna. This was made possible by the use of circular polarization, and a multi-port waveguide section operating at X band. By the late 1970s this changed to linear polarization and the use of ferrite circulators at both X and K bands. PD radars operate at too high a PRF to use a Transmit-Receive gas filled switch, and most use solid-state devices to protect the receiver Low Noise Amplifier when the transmitter is fired.

Moving target indication

Moving target indication (**MTI**) is a mode of operation of a radar to discriminate a target against clutter.

In contrast to another mode, stationary target indication, it takes an advantage of the fact that the target moves with respect to stationary clutter. The most common approach is taking an advantage of the Doppler effect. For a sequence of radar pulses the moving target will be at different distance from the radar and the phase of the radar return from the target will be different for successive pulses, while the returns from stationary clutter will arrive at the same phase shift.

Radar MTI may be specialized in terms of the type of clutter and environment: airborne MTI (**AMTI**), ground MTI (**GMTI**), etc., or may be combined mode: stationary and moving target indication (**SMTI**).

Design Parameters

Basic Geometry

A platform is going at velocity v_p with a maximum range R_{max} with elevation angle *EL* and azimuth *AZ* to the target.

Probability of Detection (Pd)

The probability of detecting a given target at a given range any time the radar beam scans across it, Pd is determined by factors that include the size of the antenna and the amount of power it radiates. A large antenna radiating at high power provides the best performance. For high quality information on moving targets the Pd must be very high.

Target Location Accuracy

Location accuracy is a function of platform self-location performance, radar-pointing accuracy, azimuth resolution, and range resolution. A long antenna or very short wave length can provide fine azimuth resolution. Short antennas tend to have a larger azimuth error, an error that increases with range to the target because signal-to-noise ratio varies inversely with range. Location accuracy is vital to tracking performance because it prevents track corruption when there are multiple targets and makes it possible to determine which road a vehicle is on if it is moving in an area with many roads.

The target location accuracy is proportional to the slant range, frequency and aperture length.

Target Range Resolution (High Range Resolution or HRR)

Target range resolution determines whether two or more targets moving in close proximity will be detected as individual targets. With higher performance radars, target range resolution—known as High Range Resolution (HRR)—can be so precise that it may be possible to recognize a specific target (i.e., one that has been seen before) and to place it in a specific class (e.g., "a T-80 tank"). This would allow more reliable tracking of specific vehicles or groups of vehicles, even when they are moving in dense traffic or disappear for a period due to screening.

Minimum Detectable Velocity MDV

The MDV comes from the frequency spread of the mainlobe clutter. MDV determines whether the majority of military traffic, which often moves very slowly, especially when traveling off-road, will be detected. A GMTI radar must distinguish a moving target from ground clutter by using the target's Doppler signature to detect the radial component of the target's velocity vector (i.e., by measuring the component of the target's movement directly along the radar-target line). To capture most of this traffic, even when it is moving almost tangentially to the radar (i.e., perpendicular to the radar-target line), a system must have the ability to detect very slow radial velocities. As the radial component of a target's velocity approaches zero, the target will fall into the clutter or *blind zone*. This is calculated as:

$$MDV = \frac{\lambda}{2}\left(\frac{4v_p}{B}\sqrt{(sin(AZ)sin(EL))^2 + (cos(AZ)cos(EL))^2}\right)$$

Any target with a velocity less than this minimum (MDV) cannot be detected because there is not sufficient Doppler shift in its echo to separate it from the mainlobe clutter return.

Area Search Rate

The area coverage rate (measured in area per unit time) is proportional to system power and aperture size. Other factors which may be relevant include grid spacing, size of the power amp, module quantization, the number of beams processed and system losses.

Stand-off Distance

Stand-off distance is the distance separating a radar system from the area it is covering.

Coverage Area Size (breadth and depth)

Coverage area size is the area that the system can keep under continuous surveillance from a specific orbit. Well known design principles cause a radar's maximum detection range to depend on the size of its antenna (radar aperture), the amount of power radiated from the antenna, and the effectiveness of its clutter cancellation mechanism. The earth's curvature and screening from terrain, foliage, and buildings cause system altitude to be another key factor determining depth of coverage. The ability to cover an area the size of an army corps commander's area of interest from a safe stand-off distance is the hallmark of an effective, advanced GMTI system.

Coverage Area Revisit Rate

This equates to the frequency with which the radar beam passes over a given area. Frequent revisits are very important to the radar's ability to achieve track continuity and contribute to an increased probability of target detection by lessening the chance of obscuration from screening by trees, buildings, or other objects. A fast revisit rate becomes critical to providing an uncorrupted track when a target moves in dense traffic or is temporarily obscured, if only by trees along a road.

Chapter-6

Pulse-Doppler Radar

Pulse-Doppler is a radar system capable of not only detecting target location (bearing, range, and altitude), but also measuring its radial velocity (range-rate). It uses the Doppler effect to determine the relative velocity of objects; pulses of RF energy returning from the target are processed to measure the phase shift between carrier cycles successive pulses at the transmitted frequency. To achieve this, the transmitter frequency source must have very good phase stability and the system is said to be coherent.

They are used in different fields. In meteorological radars, pulse-Doppler measures instantaneous speed at discrete range intervals to detect wind field in weather as the beam is slewed across the sky. Doppler On Wheels, NEXRAD, Terminal Doppler Weather Radar, and ARMOR Doppler Weather Radar are examples. In search and track radars, the purpose is to measure the speed of all targets and eliminate environmental reflections from weather, the surface of the earth, and biological objects like birds, which hides target signals, by their characteristic velocity.

Pulse repetition frequency

Pulse-Doppler typically uses medium Pulse repetition frequency (PRF) from about 3 kHz to 30 kHz. Systems using PRF below 3 kHz are considered low PRF because direct range can be measured to a range of 50 km, and Doppler processing becomes an increasing challenge due to coherency limitations as PRF falls below 3 kHz. Systems using PRF above 30 kHz function better as interrupted CW radar because direct velocity can be measured up to 4.5 km/s at L band, and range measurement becomes more of a challenge as PRF increases above 30 kHz. Range and velocity cannot be measured directly using medium PRF, and a technique called ambiguity resolutions is used to identify range.

Underlying principle

Pulse-Doppler radar is based on the Doppler effect, where movement in range produces frequency shift on the signal reflected from the target.

$$Doppler Frequency = 2F_o\left(\frac{RangeVelocity}{C}\right)$$

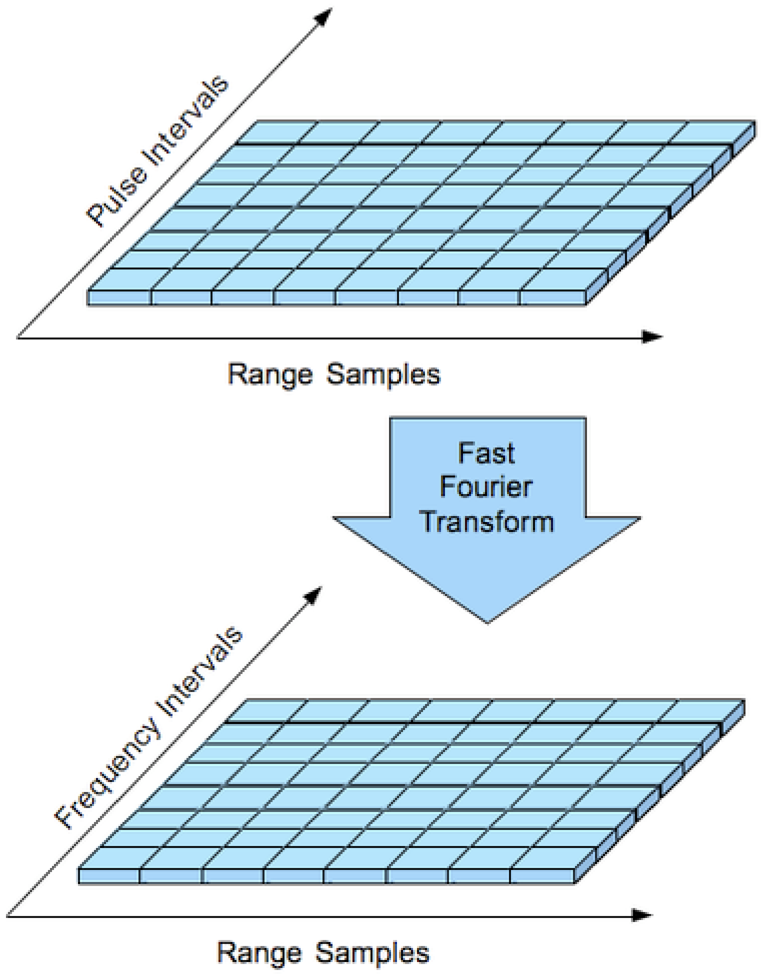
In some systems, the velocity measurement is compared to the change in range to determine if the detected signal is an airborne vehicle or an electronic artifact. Angular measurement is made using monopulse scanning and sidelobe blanking as the radar beam is scanned across the sky.

Multiple receive samples are taken between transmit pulses. Each sample includes an in-phase (I) value and a quadrature (Q) value. I an Q are the real and imaginary component of a complex number.

A spinning wheel, mirror and strobe-light can be used to visualize I and Q. The mirror is placed at a 45 degree angle above the wheel so that you can see the front and top of the wheel at the same time. The strobe-light is attached to the wheel so that you can see the wheel spin when the room lights are turned off. You sit directly in front of the wheel while a friend spins the wheel. The view of the front of the wheel (I) and the top of the wheel (Q) tell you whether your friend has spun the wheel clockwise or counterclockwise. Clockwise is like inbound Doppler. Counterclockwise is like outbound Doppler.

Pairs of I and Q values form a complex number explained as real and imaginary parts. I is the real components. Q is the imaginary component.

Signal Processing



Pulse-Doppler signal processing. The *Range Sample* axis represents individual samples taken in between each transmit pulse. The *Range Interval* axis represents each successive transmit pulse interval during which samples are taken. The Fast Fourier Transform process converts time-domain samples into frequency domain spectra.

I(t) and Q(t) are both required during the sampling process so that radar signal processing can include information about closing (approaching) versus opening (leaving) Doppler velocities. This is crucial for proper operation of pulse-Doppler systems.

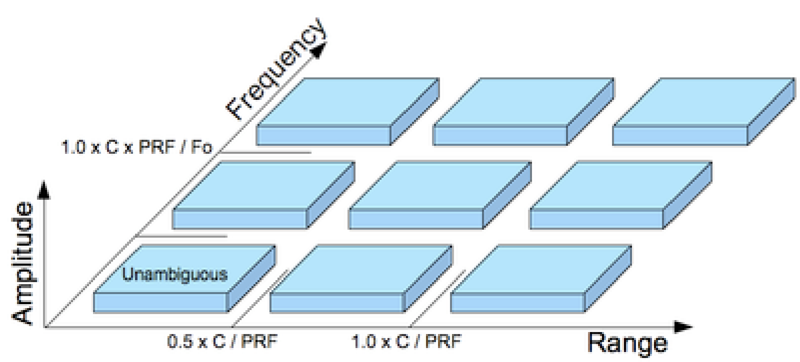
Pulse-Doppler relies on taking up to a few dozen individual I and Q samples in between each transmit pulse. Each different range sample has its own separate processing, and these time samples are converted to frequency domain using a digital filter, usually involving a fast Fourier transform (FFT). Side-lobes are produced during signal processing and a side-lobe suppression strategy, such as Dolph-Chebyshev Windowing, is commonly used to reduce false alarms.

The digital filter produces as many frequency outputs as there are time inputs. Production of one FFT requires as many samples as the number of points in the FFT (a 256 point FFT requires 256 transmit pulses). Each separate spectrum undergoes separate detection processing.

Pulse-Doppler typically uses a form of constant false alarm rate detection processing to identify target signals. Detection processing for pulse-Doppler produces an ambiguous range and ambiguous velocity corresponding to one of the FFT outputs from one of the range samples. The reflection from clutter falls into filters corresponding to a different frequency than the filter corresponding to most aircraft. Sub-clutter visibility in a pulse-Doppler radar is a measure of the power difference between the smallest signals that can be detected in the presence of a large signal at a different frequency. This Measure of Performance is called Sub-clutter Visibility . Sub-clutter visibility for pulse-Doppler is typically over 50dB. Conventional non-coherent radar using Moving target indication provides up to about 25dB sub-clutter visibility. Pulse-Doppler is used in many military applications for this reason.

Pulse-Doppler signals are audible, so a helicopter sounds like a helicopter and a jet sounds like a jet. The actual size of the target can be calculated using these audible signals, which is impossible with pulse compression and low PRF systems. Audible Doppler and target size support passive vehicle type classification when identification friend or foe is not available. This is another reason pulse-Doppler has military applications.

Ambiguity Resolution



Pulse-Doppler ambiguity zones. Each blue zone with no label represents a velocity/range combination that will be folded into the unambiguous zone. Areas outside the blue zones are blind ranges and blind velocities. Blind velocities correspond with the clutter notch.

Pulse Doppler relies on medium pulse repetition frequency (PRF) from about 3 kHz to 30 kHz. Each transmit pulse is separated by between 5 km and 50 km of distance. Operation in an environment larger than 5 km to 50 km means reflections arrive at the antenna from multiple ranges simultaneously. The true range and speed of the target is effectively folded by a modulo operation produced by the transmit and receive process.

Range Ambiguity Resolution

True range is found using two different PRF. Each pair of PRF is a separate detection scheme with unique mathematical properties that are configured into the radar. Target detection requires samples from both PRF, which are transmitted alternating back and forth rapidly during the search process.

The time between transmit pulses is split into several range samples. Each range sample contains the receive signal from a specific range interval. This is called sampling.

This is explained best using the following example, where PRF A produces a transmit pulse every 6 km and PRF B produces a transmit pulse every 5 km.

Transmit	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**
		Target PRF A			
				Target PRF B	

The apparent range for PRF A falls in the 2 km sample, and the apparent range for PRF B falls in the 4 km sample. This combination places the true target distance at 14 km (2x6+2 or 2x5+4). This can be seen graphically when range intervals are stacked end-to-end as shown below.

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 29 29

A A A A A

B B B B B B

"A" represents target range possibilities for PRF A, and "B" represents target range possibilities for PRF B.

This process is usually automated with a look-up table, and the size of this table limits the maximum range. A convolution algorithm may be used instead of a table.

Each individual PRF has blind ranges, where the transmitter pulse occurs at the same time as the target reflection signal arrives back at the radar. A third and fourth PRF are required to complete a detection by repeating the process described above. PRF alternates rapidly while scanning to fill in any blind ranges.

Multiple target reflections can corrupt the range ambiguity resolution process. Frequency ambiguity resolution improves performance for multiple target scenarios by separating target signals using speed differences.

Frequency Ambiguity Resolution

Each range samples is converted from time domain I/Q samples into frequency domain. Older systems use individual filters for this. Newer systems use digital sampling and a Fast Fourier transform or Discrete Fourier transform instead of physical filters. Each filter converts time samples into a frequency spectrum. Each spectrum frequency corresponds with a different speed. The ambiguous velocity is as follows.

$$AmbiguousVelocity = -0.5\left(\frac{DopplerFrequency * C}{TransmitFrequency}\right)$$

Frequency is folded for high speed targets where radial velocity produces a frequency shift above the Nyquist frequency. The true speed of the target may be folded by a modulo operation produced by the sampling process.

$$TrueVelocity = AmbiguousVelocity + 0.5N\left(\frac{PRF * C}{TransmitFrequency}\right)$$

$$N = IntegerBetween \pm \left(\frac{0.5Bandwidth}{PRF}\right)$$

The Nyquist frequency will also change when the PRF is changed. The target frequency will shift if the speed is greater than the Nyquist frequency. This shift can be used to establish the frequency interval, much the same way as described above for range ambiguity resolution.

A blind velocity occurs when Doppler frequency falls close to the PRF. This folds the return signal into the same filter as stationary clutter reflections. Rapidly alternating different PRF while scanning eliminates blind frequencies.

Special Consideration

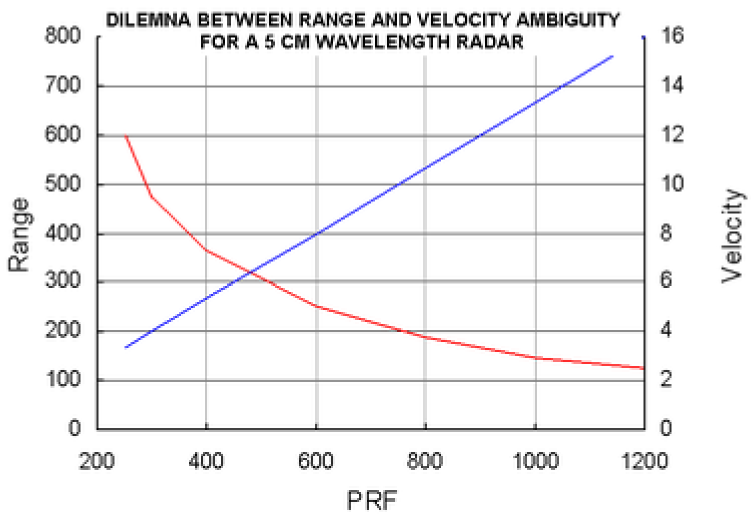
Pulse-Doppler radar has special requirements that must be satisfied to achieve acceptable performance.

Coherency

Pulse-Doppler radar requires an oscillator with very little noise. The amount of tolerable FM or phase noise depends upon the size of the filters used for signal processing. If range samples are split into 100 different frequencies using a PRF of 10 kHz, then the tolerable limit for FM transmit noise requires signal no greater than the sub-clutter visibility reduction factor 100 Hz away from the carrier frequency, and this must be maintained over a time span of no less than 10ms.

FM noise above this threshold paints stationary objects with false Doppler signal, which reduces sub-clutter visibility performance. Noise reduction includes shaping the transmit pulse to reduce ringing artifacts. Pulse Doppler requires coherent amplifiers, like solid state, klystron or traveling wave tube. Extra FM noise originating from transmitter amplification will induce apparent motion on stationary objects which will reduce performance, so magnetron and crossed-field amplifier are not appropriate because noise introduced by these devices interfere with detection performance.

In order for Pulse-Doppler radar to be effective, it is essential that the received echoes are coherent with the carrier signal, at least during the sampling period. To achieve this, a number of techniques are employed, the most common being that the transmitter signal is derived from a highly stable oscillator (the COHO) and the received signal is demodulated using an equally stable local oscillator (the STALO), which is phase locked to it. Doppler shift may then be accurately resolved by comparing the frequency components of the returned echo with the frequency components of the transmitted signal.



Maximum range from reflectivity (red) and unambiguous Doppler velocity range (blue) with a fix pulse repetition rate.

Scalloping

Scalloping is a phenomenon unique to two-PRF detection schemes. This creates **detection gaps** with a pattern of discrete ranges, each of which has a blind velocity.

Objects moving at the radial velocity that produces a harmonic of the PRF produce zero Doppler at the receiver. This is the blind velocity for a PRF. The target will become invisible in a two PRF detection scheme when the target reflection arrives during the transmit pulse on the other PRF.

Assume two pulse frequencies PRF-A and PRF-B. There will be blind zones at every range for PRF-A where there is a transmit pulse at the same time as when a signal is arriving. PRF-B will produce detections in those blind ranges, except that PRF-B will have blind velocities different from PRF-A.

This leaves intermittent detection gaps at discrete velocities in each blind range. Aircraft with a Doppler frequency that is a harmonic of PRF-B will not be detectable at blind ranges for PRF-A. Aircraft with a Doppler frequency that is a harmonic of PRF-A will not be detectable at blind ranges for PRF-B.

These detection gaps are filled in by using three or more alternating PRF in the detection scheme. The other alternative is to increase the pulse rate high enough to eliminate frequency ambiguity.

Velocity becomes ambiguous at a 3 kHz PRF for an L band radar at 300 m/s (near the speed of sound). Velocity becomes ambiguous at a 3 kHz PRF for an X band radar at 30 m/s (speed of a fast car).

Increasing PRF increases the Nyqist frequency. Velocity becomes ambiguous at a 30 kHz PRF for an L band radar at 3000 m/s (about half the velocity of low earth orbit). Velocity becomes ambiguous at a 30 kHz PRF for an X band radar at 300 m/s (speed of sound).

Alternating PRF requires a minimum level of complexity that depends upon maximum radial velocity of the object. The maximum radial velocity is determined by the environment. Objects moving inside earth's atmosphere are limited to about 1 km/s due to thermal heating caused by shock wave physics. Space begins above about 20miles, and objects in low earth orbit travel about 8 km/s. There is no velocity limit in space.

Modeling and simulation is required used to evaluate performance for the chosen PRF combination to ensure no detection gaps for all possible range and velocity combination.

Range Performance

The theoretical range performance is as follows.

$$R = \left(\frac{P_t G_t A_r \sigma F^4 D}{16\pi^2 K_b T B N} \right)^{-4}$$

where

- R = distance to the target
- P_t = transmitter power
- G_t = gain of the transmitting antenna
- A_r = effective aperture (area) of the receiving antenna
- σ = radar cross section, or scattering coefficient, of the target

- F = pattern propagation factor
- D = number of discrete Doppler filters for each range sample
- R = distance to the target
- K_b = Boltzmann's constant
- T = Temperature (kelvin)
- B = Receiver Bandwidth (band pass filter)
- N = Noise Factor

This is derived by combining the Radar equation with the Noise equation and accounting for in-band noise distribution across multiple detection filters. The value D is added to the standard radar range equation to account for noise power that arrives at the receiver being distributed equally among each of the Doppler filters.

Angular movement of the antenna must be slow enough so each volume of space remains within the main antenna lobe no less than the duration of three full PRF samples to achieve the level of performance described by this equation. Increasing the value of D will increase detection range at the expense of requiring more time to complete a full scan. It is impractical to increase D to a value that that requires phase noise lower than the transmitter specification.

Detection range is increased by the number of filters. For example, use of 100 filters will reduce noise contribution in each filter 20dB below the level of noise (electronics) begin sampled in the receiver. Each filter holds a small amount of the total noise arriving at the receiver but all of the signal arriving from the target goes into no more than two filters. This reduces the bandwidth associated with the detection process. A 20dB improvement corresponds with the ability to detect a target 100 times smaller than with conventional pulse-amplitude radar, all else being equal.

Active RF phase shifters commonly used with electronically steered phased array antenna normally have a phase settling time that exceeds the coherency limit needed for pulse-Doppler. This limits pulse-Doppler antennas to a mechanically steered configuration.

Application considerations

Type of Radar

The maximum velocity that can be unambiguously measured is inherently limited by the PRF, as discussed above. The PRF-value must therefore be chosen carefully, based on a tradeoff between maximum velocity resolution and the reduction of velocity aliasing and range ambiguity problems. This tradeoff is highly application dependent, as e.g. weather radars measure velocities at a totally different scale as compared to radars designed to detect supersonic missiles and aircraft.

Moving targets

Stationary targets such as earth ground clutter (land, buildings, etc.) will be dominant in the low doppler frequencies, while moving targets will produce much higher doppler shifts. The radar processor can be designed to mask out clutter by the use of doppler filters (digital or analogue) around the main spectral line (called the clutter-notch), which will result in the display of moving targets only (in relation to the radar). If the radar itself is moving, such as on a fighter aircraft, or a surveillance aircraft, then much more processing will be required, as the clutter in the filters will be based on platform speed, terrain under the radar, antenna depression angle, and antenna rotation/steered angle.

Monopulse radar

Doppler signal processing improves performance of monopulse radar systems by reducing or eliminating signals from slow object, like the earth surface, trees, buildings, and weather.

Chapter-7

Radar Tracker

A **radar tracker** is a component of a radar system, or an associated command and control (C2) system, that associates consecutive radar observations of the same target into tracks. It is particularly useful when the radar system is reporting data from several different targets or when it is necessary to combine the data from several different radars or other sensors.

Role of the radar tracker

A classical rotating air surveillance radar system detects target echoes against a background of noise. It reports these detections (known as "plots") in polar coordinates representing the range and bearing of the target. In addition, noise in the radar receiver will occasionally exceed the detection threshold of the radar's Constant false alarm rate detector and be incorrectly reported as targets (known as false alarms). The role of the radar tracker is to monitor consecutive updates from the radar system (which typically occur once every few seconds, as the antenna rotates) and to determine those sequences of plots belonging to the same target, whilst rejecting any plots believed to be false alarms. In addition, the radar tracker is able to use the sequence of plots to estimate the current speed and heading of the target. When several targets are present, the radar tracker aims to provide one track for each target, with the track history often being used to indicate where the target has come from.

When multiple radar systems are connected to a single reporting post, a **multiradar tracker** is often used to monitor the updates from all of the radars and form tracks from the combination of detections. In this configuration, the tracks are often more accurate than those formed from single radars, as a greater number of detections can be used to estimate the tracks. In addition to associating plots, rejecting false alarms and estimating heading and speed, the radar tracker also acts as a filter, in which errors in the individual radar measurements are smoothed out. In essence, the radar tracker fits a smooth curve to the reported plots and, if done correctly, can increase the overall accuracy of the radar system. A **multisensor tracker** extends the concept of the multiradar tracker to allow the combination of reports from different types of sensor - typically radars, secondary surveillance radars, identification friend or foe (IFF) systems and electronic support measures (ESM) data. A radar track will typically contain the following information

- Position (in two or three dimensions)

- Heading
- Speed
- Unique track number

In addition, and depending on the application or tracker sophistication, the track will also include:

- Civilian SSR Modes A, C, S information
- Military IFF Modes 1, 2, 3, 4 and 5 information
- Call sign information
- ADS-B information
- Track reliability or uncertainty information

General approach

There are many different mathematical algorithms used for implementing a radar tracker, of varying levels of sophistication. However, they all perform steps similar to the following every time the radar updates:

- Associate a radar plot with an existing track (plot to track association)
- Update the track with this latest plot (track smoothing)
- Spawn new tracks with any plots that are not associated with existing tracks (track initiation)
- Delete any tracks that have not been updated, or predict their new location based on the previous heading and speed (track maintenance)

Perhaps the most important step is the updating of tracks with new plots. All trackers will implicitly or explicitly take account of a number of factors during this stage, including:

- a model for how the radar measurements are related to the target coordinates
- the errors on the radar measurements
- a model of the target movement
- errors in the model of the target movement

Using these information, the radar tracker attempts to update the track by forming a weighted average of the current reported position from the radar (which has unknown errors) and the last predicted position of the target from the tracker (which also has unknown errors). The tracking problem is made particularly difficult for targets with unpredictable movements (i.e. unknown target movement models), non-Gaussian measurement or model errors, non-linear relationships between the measured quantities and the desired target coordinates, detection in the presence of non-uniformly distributed clutter, missed detections or false alarms. In the real world, a radar tracker typically faces a combination of all of these effects; this has led to the development of an increasingly sophisticated set of algorithms to resolve the problem. Due to the need to form radar tracks in real time, usually for several hundred targets at once, the deployment of radar tracking algorithms has typically been limited by the available computational power.

Plot to track association

In this step of the processing, the radar tracker seeks to determine which plots should be used to update which tracks. In many approaches, a given plot can only be used to update one track. However, in other approaches a plot can be used to update several tracks, recognising the uncertainty in knowing to which track the plot belongs. Either way, the first step in the process is to update all of the existing tracks to the current time by predicting their new position based on the most recent state estimate (e.g. position, heading, speed, acceleration, etc.) and the assumed target motion model (e.g. constant velocity, constant acceleration, etc.). Having updated the estimates, it is possible to try to associate the plots to tracks.

This can be done in a number of ways:

- By defining an "acceptance gate" around the current track location and then selecting:
    - the closest plot in the gate to the predicted position, or
    - the strongest plot in the gate
- By a statistical approach, such as the Probabilistic Data Association Filter (PDAF) or the Joint Probabilistic Data Association Filter (JPDAF) that choose the most probable location of plot through a statistical combination of all the likely plots. This approach has been shown to be good in situations of high radar clutter.

Once a track has been associated with a plot, it moves to the **track smoothing** stage, where the track prediction and associated plot are combined to provide a new, smoothed estimate of the target location.

Having completed this process, a number of plots will remain unassociated to existing tracks and a number of tracks will remain without updates. This leads to the steps of **track initiation** and **track maintenance**.

Track initiation

Track initiation is the process of creating a new radar track from an unassociated radar plot. Obviously, when the tracker is first switched on, all of the initial radar plots are used to create new tracks, but once the tracker is running, only those plots that couldn't be used to update an existing track are used to spawn new tracks. Typically a new track is given the status of **tentative** until plots from subsequent radar updates have been successfully associated with the new track. Tentative tracks are not shown to the operator and so they provide a means of preventing false tracks from appearing on the screen - at the expense of some delay in the first reporting of a track. Once several updates have been received, the track is **confirmed** and displayed to the operator. The most common criterion for promoting a tentative track to a confirmed track is the "M-of-N rule", which states that during the last N radar updates, at least M plots must have been associated with the tentative track - with M=3 and N=5 being typical values. More sophisticated approaches may use a statistical approach in which a track becomes confirmed when, for instance, its covariance matrix falls to a given size.

Track maintenance

Track maintenance is the process in which a decision is made about whether to end the life of a track. If a track was not associated with a plot during the plot to track association phase, then there is a chance that the target may no longer exist (for instance, an aircraft may have landed or flown out of radar cover). Alternatively, however, there is a chance that the radar may have just failed to see the target at that update, but will find it again on the next update. Common approaches to deciding on whether to terminate a track include:

- If the target was not seen for the past M consecutive update opportunities (typically M=3 or so)

- If the target was not seen for the past M out of N most recent update opportunities

- If the target's track uncertainty (covariance matrix) has grown beyond a certain threshold

Track smoothing

In this important step, the latest track prediction is combined with the associated plot to provide a new, improved estimate of the target state as well as a revised estimate of the errors in this prediction. There are a wide variety of algorithms, of differing complexity and computational load, that can be used for this process.

Alpha-beta tracker

An early tracking approach that assumed fixed Gaussian errors and a constant-speed, non-maneuvering target model to update tracks.

Kalman filter

The role of the Kalman Filter is to take the current known state (i.e. position, heading, speed and possibly acceleration) of the target and predict the new state of the target at the time of the most recent radar measurement. In making this prediction, it also updates its estimate of its own uncertainty (i.e. errors) in this prediction. It then forms a weighted average of this prediction of state and the latest measurement of state, taking account of the known measurement errors of the radar and its own uncertainty in the target motion models. Finally, it updates its estimate of its uncertainty of the state estimate. A key assumption in the mathematics of the Kalman filter is that measurement equations (i.e. the relationship between the radar measurements and the target state) and the state equations (i.e. the equations for predicting a future state based on the current state) are linear - i.e. can be expressed in the form $y = A.x$ (where A is a constant), rather than $y = f(x)$.

The Kalman filter assumes that the measurement errors of the radar, and the errors in its target motion model, and the errors in its state estimate are all zero-mean Gaussian distributed. This means that all of these sources of errors can be represented by a covariance matrix. The mathematics of the Kalman filter is therefore concerned with propagating these covariance matrices and using them to form the weighted sum of prediction and measurement.

In situations where the target motion conforms well to the underlying model, there is a tendency of the Kalman filter to become "over confident" of its own predictions and to start to ignore the radar measurements. If the target then manoeuvres, the filter will fail to follow the manoeuvre. It is therefore common practice when implementing the filter to arbitrarily increase the magnitude of the state estimate covariance matrix slightly at each update to prevent this.

Multiple hypothesis tracker (MHT)

The MHT allows a track to be updated by more than one plot at each update, spawning multiple possible tracks. As each radar update is received every possible track can be potentially updated with every new update. Over time, the track branches into many possible directions. The MHT calculates the probability of each potential track and typically only reports the most probable of all the tracks. For reasons of finite computer memory and computational power, the MHT typically includes some approach for deleting the most unlikely potential track updates. The MHT is designed for situations in which the target motion model is very unpredictable, as all potential track updates are considered. For this reason, it is popular for problems of ground target tracking in Airborne Ground Surveillance (AGS) systems.

Interacting multiple model (IMM)

The IMM is an estimator which can either be used by MHT or JPDAF. IMM uses two or more Kalman filters which run in parallel, each using a different model for target motion or errors. The IMM forms an optimal weighted sum of the output of all the filters and is able to rapidly adjust to target maneuvers. While MHT or JPDAF handles the association and track maintenance, an IMM helps MHT or JPDAF in obtaining a filtered estimate of the target position.

Nonlinear tracking algorithms

Non-linear tracking algorithms use a Non-linear filter to cope with the situation where the measurements have a non-linear relationship to the final track coordinates, where the errors are non-Gaussian, or where the motion update model is non-linear. The most common non-linear filters are:

- the Extended Kalman filter
- the Unscented Kalman filter
- the Particle filter

Extended Kalman filter (EKF)

The EKF is an extension of the Kalman filter to cope with cases where the relationship between the radar measurements and the track coordinates, or the track coordinates and the motion model, is non-linear. In this case, the relationship between the measurements and the state is of the form $h = f(x)$ (where h is the vector of measurements, x is the target state and f(.) is the function relating the two). Similarly, the relationship between the future state and the current state is of the form $x(t+1) = g(x(t))$ (where x(t) is the state at time t and g(.) is the function that predicts the future state). To handle these non-linearities, the EKF linearises the two non-linear equations using the first term of the Taylor series and then treats the problem as the standard linear Kalman filter problem. Although conceptually simple, the filter can easily diverge (i.e. gradually perform more and more badly) if the state estimate about which the equations are linearised is poor.

The unscented Kalman filter and particle filters are attempts to overcome the problem of linearising the equations.

Unscented Kalman filter (UKF)

The UKF attempts to improve on the EKF by removing the need to linearise the measurement and state equations. Although it retains the assumption that the filter errors are Gaussian distributed, rather than model these as covariance matrices, it instead chooses an explicit sample of those Gaussian errors by choosing a small number (typically 5 or so) of different state estimates that have the required mean and variance. These points are then propagated directly through the non-linear equations, and the resulting five updated samples are then used to calculate a new mean and variance. This approach then suffers none of the problems of divergence due to poor linearisation and yet retains the overall computational simplicity of the EKF.

Particle filter

The Particle Filter could be considered as a generalisation of the UKF. It makes no assumptions about the distributions of the errors in the filter and neither does it require the equations to be linear. Instead it generates a large number of random potential states ("particles") and then propagates this "cloud of particles" through the equations, resulting in a different distribution of particles at the output. The resulting distribution of particles can then be used to calculate a mean or variance, or whatever other statistical measure is required. The resulting statistics are used to generate the random sample of particles for the next iteration. The particle filter is notable in its ability to handle multi-modal distributions (i.e. distributions where the PDF has more than one peak). However, it is computationally very intensive and is currently unsuitable for most real-world, real-time applications.

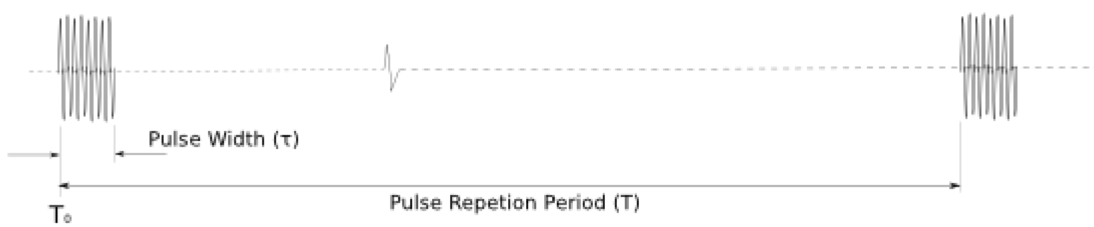
Chapter-8

Radar Signal Characteristics

A radar system uses a radio frequency electromagnetic signal reflected from a target to determine information about that target. In any radar system, the signal transmitted and received will exhibit many of the characteristics described below.

The radar signal in the time domain

The diagram below shows the characteristics of the transmitted signal in the time domain. Note that in this and in all the diagrams within here, the x axis is exaggerated to make the explanation clearer.



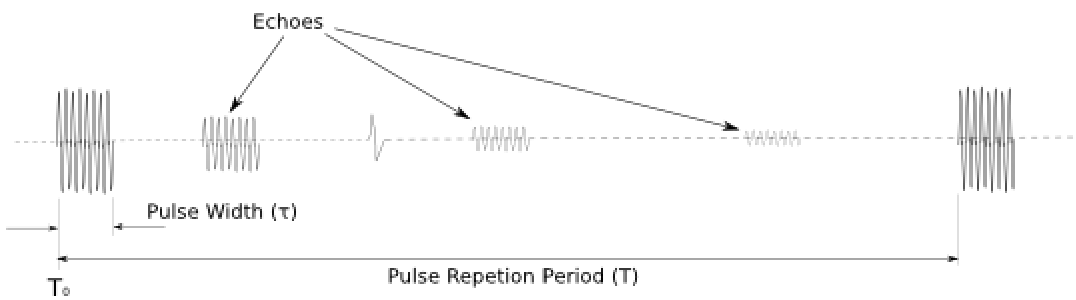
Carrier

The carrier is an RF signal, typically of microwave frequencies, which is usually (but not always) modulated to allow the system to capture the required data. In simple ranging radars, the carrier will be pulse modulated and in continuous wave systems, such as Doppler radar, modulation may not be required. Most systems use pulse modulation, with or without other supplementary modulating signals. Note that with pulse modulation, the carrier is simply switched on and off in sync with the pulses; the modulating waveform does not actually exist in the transmitted signal and the envelope of the pulse waveform is extracted from the demodulated carrier in the receiver. Although obvious when described, this point is often missed when pulse transmissions are first studied, leading to misunderstandings about the nature of the signal.

Pulse width

The pulse width (τ) (or pulse duration) of the transmitted signal is to ensure that the radar emits sufficient energy to allow that the reflected pulse is detectable by its receiver. The amount of energy that can be delivered to a distant target is the product of two things; the output power of the transmitter, and the duration of the transmission. Therefore pulse width constrains the maximum detection range of a target.

It also determines the range discrimination, that is the capacity of the radar to distinguish between two targets fairly close together. At *any* range, with similar azimuth and elevation angles and as viewed by a radar with an unmodulated pulse, the range discrimination is approximately equal in distance to half of the pulse duration.



Pulse width also determines the dead zone at close ranges. While the radar transmitter is active, the receiver input is blanked to avoid the amplifiers being swamped (saturated) or, (more likely), damaged. A simple calculation reveals that a radar echo will take approximately 10.8 μs to return from a target 1 standard mile away (counting from the leading edge of the transmitter pulse (T_0), (sometimes known as transmitter main bang)). For convenience, these figures may also be expressed as 1 nautical mile in 12.4 μs or 1 kilometre in 6.7 μs. (For simplicity, all further discussion will use metric figures.) If the radar pulse width is 1 μs, then there can be no detection of targets closer than about 150 m, because the receiver is blanked.

All this means that the designer cannot simply increase the pulse width to get greater range without having an impact on other performance factors. As with everything else in a radar system, compromises have to be made to a radar system's design to provide the optimal performance for it's role.

Pulse repetition frequency (PRF)

In order to build up a discernible echo, most radar systems emit pulses continuously and the repetition rate of these pulses is determined by the role of the system. An echo from a target will therefore be 'painted' on the display or integrated within the signal processor every time a new pulse is transmitted, reinforcing the return and making detection easier. The higher the PRF that is used, then the more the target is painted. However with the higher PRF the range that the radar can "see" is reduced. Radar designers try to use the highest PRF possible commensurate with the other factors that constrain it, as described below.

There are two other facets related to PRF that the designer must weigh very carefully; the beamwidth characteristics of the antenna, and the required periodicity with which the radar must sweep the field of view. A radar with a 1° horizontal beamwidth that sweeps the entire 360° horizon every 2 seconds with a PRF of 1080 Hz will radiate 3 pulses over each 1-degree arc. If the receiver needs at least 6 reflected pulses of similar amplitudes to achieve an acceptable probability of detection, then there are three choices for the designer: double the PRF, halve the sweep speed, or double the beamwidth. In reality, all three choices are used, to varying extents; radar design is all about compromises between conflicting pressures.

Staggered PRF

Staggered PRF is a transmission process where the time between interrogations from radar changes slightly, *in a patterned and readily-discernible repeating manner.* The change of repetition frequency allows the radar, on a pulse-to-pulse basis, to differentiate between returns from its own transmissions and returns from other radar systems with the same PRF and a similar radio frequency. Consider a radar with a constant interval between pulses; target reflections appear at a relatively constant range related to the flight-time of the pulse. In today's very crowded radio spectrum, there may be many other pulses detected by the receiver, either directly from the transmitter or as reflections from elsewhere. Because their apparent "distance" is defined by measuring their time relative to the last pulse transmitted by "our" radar, these "jamming" pulses could appear at any apparent distance. When the PRF of the "jamming" radar is very similar to "our" radar, those apparent distances may be very slow-changing, just like real targets. By using stagger, a radar designer can force the "jamming" to jump around erratically in apparent range, inhibiting integration and reducing or even suppressing its impact on true target detection.

Without staggered PRF, any pulses originating from another radar on the same radio frequency might appear stable in time and could be mistaken for reflections from the radar's own transmission. With staggered PRF the radar’s own targets appear stable in range in relation to the transmit pulse, whilst the 'jamming' echoes may move around in apparent range (uncorrelated), causing them to be rejected by the receiver. Staggered PRF is only one of several similar techniques used for this, including jittered PRF (where the pulse timing is varied in a less-predictable manner), pulse-frequency modulation, and several other similar techniques whose principal purpose is to reduce the probability of unintentional synchronicity. These techniques are in widespread use in marine safety and navigation radars, by far the most numerous radars on planet Earth today.

Clutter

Clutter (also termed *ground clutter*) is a form of radar signal contamination. It occurs when fixed objects close to the transmitter—such as buildings, trees, or terrain (hills, ocean swells and waves)—obstruct a radar beam and produce echoes. The echoes resulting from ground clutter may be large in both areal size and intensity. The effects of ground clutter fall off as range increases usually due to the curvature of the earth and the tilt of the antenna above the horizon. Without special processing techniques, targets can be lost in returns from terrain on land or waves at sea.

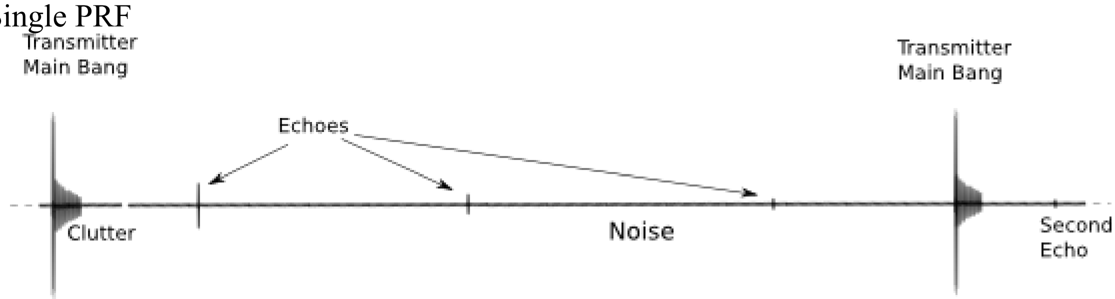
Clutter is used by the military to jam radars by the use chaff. Chaff is small reflective material used to hide troop, ship, or aircraft movements by creating many returns and overwhelming the radar's receiver with spurious targets.

Sensitivity time control (STC)

STC is used to avoid saturation of the receiver from close in ground clutter by adjusting the attenuation of the receiver as a function of distance. More attenuation is applied to returns close in and is reduced as the range increases.

Unambiguous range

Single PRF



In simple systems, echoes from targets must be detected and processed before the next transmitter pulse is generated if range ambiguity is to be avoided. Range ambiguity occurs when the time taken for an echo to return from a target is greater than the pulse repetition period (T); if the interval between transmitted pulses is 1000 microseconds, and the return-time of a pulse from a distant target is 1200 microseconds, the apparent distance of the target is only 200 microseconds. In sum, these 'second echoes' appear on the display to be targets closer than they really are.

Consider the following example : if the radar antenna is located at around 15 m above sea level, then the distance to the horizon is pretty close, (perhaps 15 km). Ground targets further than this range cannot be detected, so the PRF can be quite high; a radar with a PRF of 7.5 kHz will return ambiguous echoes from targets at about 20 km, or over the horizon. If however, the PRF was doubled to 15 kHz, then the ambiguous range is reduced to 10 km and targets beyond this range would only appear on the display after the transmitter has emitted another pulse. A target at 12 km would appear to be 2 km away, although the strength of the echo might be much lower than that from a genuine target at 2 km.

The maximum non ambiguous range varies inversely with PRF and is given by:

$$Range_{maxunambiguous} = \left(\frac{c}{2\,PRF}\right)$$

If a longer unambiguous range is required with this simple system, then lower PRFs are required and it was quite common for early search radars to have PRFs as low as a few hundred Hz, giving an unambiguous range out to well in excess of 150 km. However, lower PRFs introduce other problems, including poorer target painting and velocity ambiguity in Pulse-Doppler systems.

Multiple PRF

Modern radars, especially air-to-air combat radars in military aircraft, may use PRFs in the tens-to-hundreds of kilohertz and stagger the interval between pulses to allow the correct range to be determined. With this form of staggered PRF, a *packet* of pulses is transmitted with a fixed interval between each pulse, and then another *packet* is transmitted with a slightly different interval. Target reflections appear at different ranges for each *packet*; these differences are accumulated and then simple arithmetical techniques may be applied to determine true range. Such radars may use repetitive patterns of *packets*, or more adaptable *packets* that respond to apparent target behaviors. Regardless, radars that employ the technique are universally coherent, with a very stable radio frequency, and the pulse *packets* may also be used to make measurements of the Doppler shift (a velocity-dependent modification of the apparent radio frequency), especially when the PRFs are in the hundreds-of-kilohertz range. Radars exploiting Doppler effects in this manner typically determine relative velocity first, from the Doppler effect, and then use other techniques to derive target distance.

Maximum Unambiguous Range

At its most simplistic, MUR (Maximum Unambiguous Range) for a Pulse Stagger sequence may be calculated using the TSP (Total Sequence Period). TSP is defined as the total time it takes for the Pulsed pattern to repeat. This can be found by the addition of all the elements in the stagger sequence. The formula is derived from the speed of light and the length of the sequence:

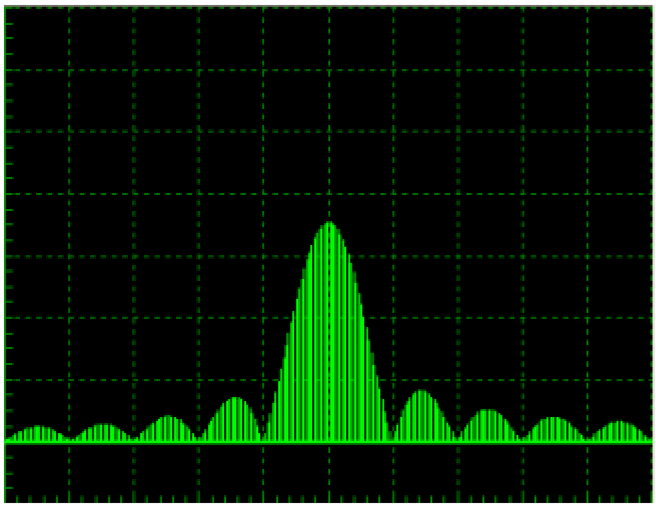
$$MUR = (c * 0.5 * TSP)$$

where c is the speed of light, usually in metres per microsecond, and TSP is the addition of all the positions of the stagger sequence, usually in microseconds. However, it should be noted that in a stagger sequence, some intervals may be repeated several times; when this occurs, it is more appropriate to consider TSP as the addition of all the **unique** intervals in the sequence.

Also, it is worth remembering that there may be vast differences between the MUR and the maximum range (the range beyond which reflections will probably be too weak to be detected), and that the maximum *instrumented* range may be *much* shorter than either of these. A civil marine radar, for instance, may have user-selectable maximum *instrumented* display ranges of 72, or 96 or rarely 120 nautical miles, in accordance with international law, but maximum unambiguous ranges of over 40,000 nautical miles and maximum detection ranges of perhaps 150 nautical miles. When such huge disparities are noted, it reveals that the primary purpose of staggered PRF is to reduce "jamming", rather than to increase unambiguous range capabilities.

The radar signal in the frequency domain

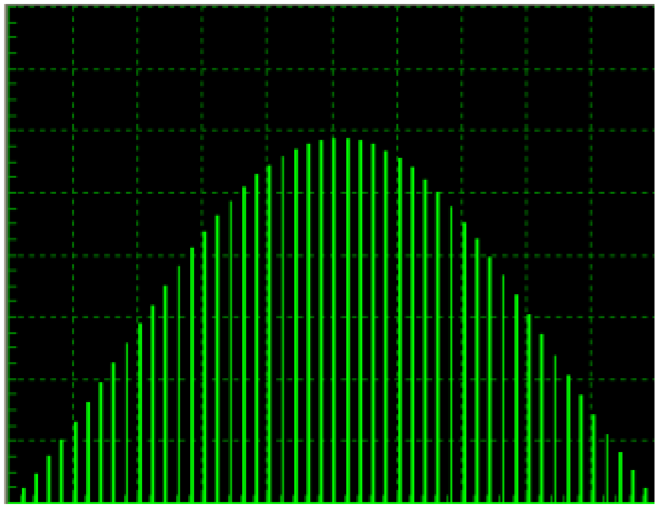
Pure CW radars appear as a single line on a Spectrum analyser display and when modulated with other sinusoidal signals, the spectrum differs little from that obtained with standard analogue modulation schemes used in communications systems, such as Frequency Modulation and consist of the carrier plus a relatively small number of sidebands. When the radar signal is modulated with a pulse train as shown above, the spectrum becomes much more complicated and far more difficult to visualise.



Basic Fourier analysis shows that any repetitive complex signal consists of a number of harmonically related sine waves. The radar pulse train is a form of square wave, the pure form of which consists of the fundamental plus all of the odd harmonics. The exact composition of the pulse train will depend on the pulse width and PRF, but mathematical analysis can be used to calculate all of the frequencies in the spectrum. When the pulse train is used to modulate a radar carrier, the typical spectrum shown on the left will be obtained.

Examination of this spectral response shows that it contains two basic structures. The Coarse Structure; (the peaks or 'lobes' in the diagram on the left) and the Fine Structure which contains the individual frequency components as shown below. The Envelope of the lobes in the Coarse Structure is given by: $\frac{1}{\pi f \tau}$.

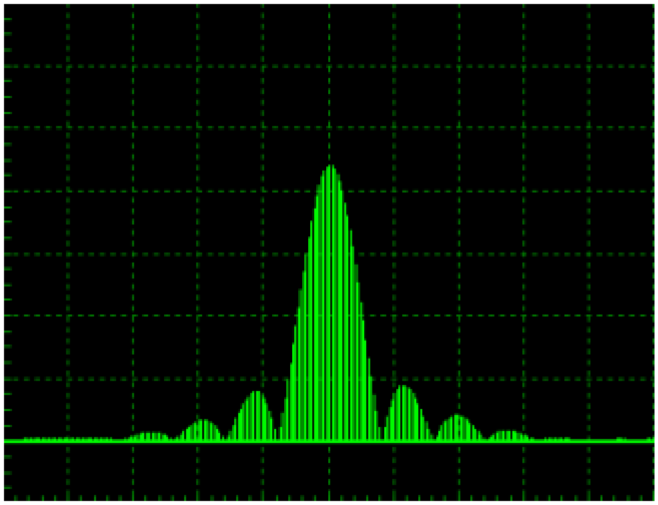
Note that the pulse width (τ) appears on the bottom of this equation and determines the lobe spacing. Smaller pulse widths result in wider lobes and therefore greater bandwidth.



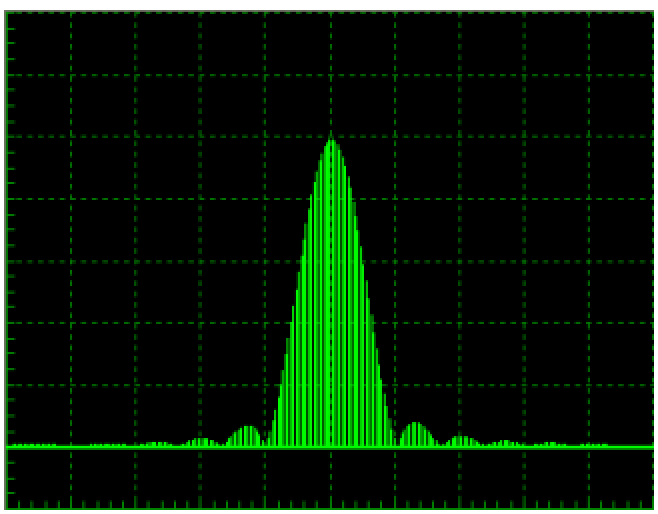
Examination of the spectral response in finer detail, as shown on the right, shows that the Fine Structure contains individual lines or spot frequencies. The formula for the fine structure is given by $\frac{T}{\pi f \tau}$ and since the period of the PRF (T) appears at the top of the fine spectrum equation, there will be fewer lines if higher PRFs are used. These facts affect the decisions made by radar designers when considering the trade-offs that need to be made when trying to overcome the ambiguities that affect radar signals.

Pulse profiling

If the rise and fall times of the modulation pulses are zero, (e.g. the pulse edges are infinitely sharp), then the sidebands will be as shown in the spectral diagrams above. The bandwidth consumed by this transmission can be huge and the total power transmitted is distributed over many hundreds of spectral lines. This is a potential source of interference with any other device and frequency-dependent imperfections in the transmit chain mean that some of this power never arrives at the antenna. In reality of course, it is impossible to achieve such sharp edges, so in practical systems the sidebands contain far fewer lines than a perfect system. If the bandwidth can be limited to include relatively few sidebands, by rolling off the pulse edges intentionally, an efficient system can be realised with the minimum of potential for interference with nearby equipment. However, the trade-off of this is that slow edges make range resolution poor. Early radars limited the bandwidth through filtration in the transmit chain, e.g. the waveguide, scanner etc., but performance could be sporadic with unwanted signals breaking through at remote frequencies and the edges of the recovered pulse being indeterminate. Further examination of the basic Radar Spectrum shown above shows that the information in the various lobes of the Coarse Spectrum is identical to that contained in the main lobe, so limiting the transmit and receive bandwidth to that extent provides significant benefits in terms of efficiency and noise reduction.



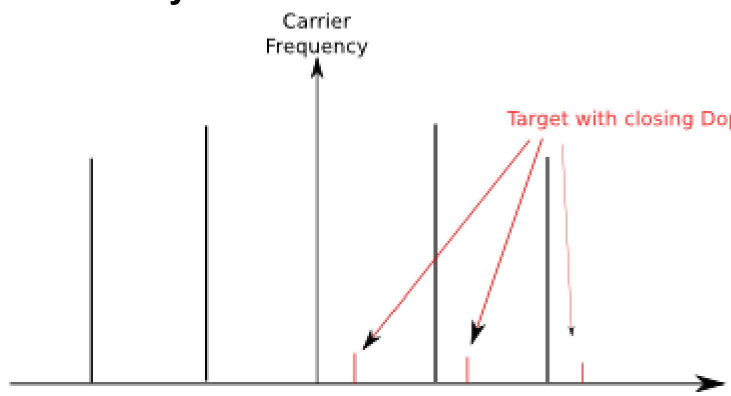
Recent advances in signal processing techniques have made the use of pulse profiling or shaping more common. By shaping the pulse envelope before it is applied to the transmitting device, say to a cosine law or a trapezoid, the bandwidth can be limited at source, with less reliance on filtering. When this technique is combined with pulse compression, then a good compromise between efficiency, performance and range resolution can be realised. The diagram on the left shows the effect on the spectrum if a trapezoid pulse profile is adopted. It can be seen that the energy in the sidebands is significantly reduced compared to the main lobe and the amplitude of the main lobe is increased.



Similarly, the use of a cosine pulse profile has an even more marked effect, with the amplitude of the sidelobes practically becoming negligible. The main lobe is again increased in amplitude and the sidelobes correspondingly reduced, giving a significant improvement in performance.

There are many other profiles that can be adopted to optimise the performance of the system, but cosine and trapezoid profiles generally provide a good compromise between efficiency and resolution and so tend to be used most frequently.

Unambiguous velocity



This is an issue only with a particular type of system; the Pulse-Doppler radar, which uses the Doppler effect to resolve velocity from the apparent change in frequency caused by targets that have net radial velocities compared to the radar device. Examination of the spectrum generated by a pulsed transmitter, shown above, reveals that each of the sidebands, (both coarse and fine), will be subject to the Doppler effect, another good reason to limit bandwidth and spectral complexity by pulse profiling.

Consider the positive shift caused by the closing target in the diagram which has been highly simplified for clarity. It can be seen that as the relative velocity increases, a point will be reached where the spectral lines that constitute the echoes are hidden or aliased by the next sideband of the modulated carrier. Transmission of multiple pulse-packets with different PRF-values, e.g. staggered PRFs, will resolve this ambiguity, since each new PRF value will result in a new sideband position, revealing the velocity to the receiver. The maximum unambiguous target velocity is given by:

$$\pm\frac{cPRF}{4f}$$

Typical system parameters

Taking all of the above characteristics into account means that certain constraints are placed on the radar designer. For example, a system with a 3 GHz carrier frequency and a pulse width of 1us will have a carrier period of approximately 333ps. Each transmitted pulse will contain about 3000 carrier cycles and the velocity and range ambiguity values for such a system would be:

PRF	Velocity Ambiguity	Range Ambiguity
Low (2 kHz)	50 m/s	75 km
Medium (12 kHz)	300 m/s	12.5 km
High (200 kHz)	5000 m/s	750 m

Chapter-9

Johnson–Nyquist Noise

Johnson–Nyquist noise (**thermal noise**, **Johnson noise**, or **Nyquist noise**) is the electronic noise generated by the thermal agitation of the charge carriers (usually the electrons) inside an electrical conductor at equilibrium, which happens regardless of any applied voltage.

Thermal noise is approximately white, meaning that the power spectral density is nearly equal throughout the frequency spectrum. Additionally, the amplitude of the signal has very nearly a Gaussian probability density function.

History

This type of noise was first measured by John B. Johnson at Bell Labs in 1928. He described his findings to Harry Nyquist, also at Bell Labs, who was able to explain the results.

Noise voltage and power

Thermal noise is distinct from shot noise, which consists of additional current fluctuations that occur when a voltage is applied and a macroscopic current starts to flow. For the general case, the above definition applies to charge carriers in any type of conducting medium (e.g. ions in an electrolyte), not just resistors. It can be modeled by a voltage source representing the noise of the non-ideal resistor in series with an ideal noise free resistor.

The one-sided power spectral density, or voltage variance (mean square) per hertz of bandwidth, is given by

$$\bar{v_n^2} = 4k_BTR$$

where k_B is Boltzmann's constant in joules per kelvin, T is the resistor's absolute temperature in kelvins, and R is the resistor value in ohms (Ω). Use this equation for quick calculation:

$$\sqrt{\bar{v_n^2}} = 0.13\sqrt{R}\ \mathrm{nV}/\sqrt{\mathrm{Hz}}.$$

For example, a 1 kΩ resistor at a temperature of 300 K has

$$\sqrt{\bar{v_n^2}} = \sqrt{4 \cdot 1.38 \cdot 10^{-23}\ \mathrm{J/K} \cdot 300\ \mathrm{K} \cdot 1\ \mathrm{k\Omega}} = 4.07\ \mathrm{nV}/\sqrt{\mathrm{Hz}}.$$

For a given bandwidth, the root mean square (RMS) of the voltage, v_n, is given by

$$v_n = \sqrt{\bar{v_n^2}}\sqrt{\Delta f} = \sqrt{4k_B T R \Delta f}$$

where Δf is the bandwidth in hertz over which the noise is measured. For a 1 kΩ resistor at room temperature and a 10 kHz bandwidth, the RMS noise voltage is 400 nV. A useful rule of thumb to remember is that 50 Ω at 1 Hz bandwidth correspond to 1 nV noise at room temperature.

A resistor in a short circuit dissipates a noise power of

$$P = \bar{v_n^2}/R = 4k_B T \Delta f.$$

The noise generated at the resistor can transfer to the remaining circuit; the maximum noise power transfer happens with impedance matching when the Thévenin equivalent resistance of the remaining circuit is equal to the noise generating resistance. In this case each one of the two participating resistors dissipates noise in both itself and in the other resistor. Since only half of the source voltage drops across any one of these resistors, the resulting noise power is given by

$$P = k_B T \Delta f$$

where P is the thermal noise power in watts. Notice that this is independent of the noise generating resistance.

Noise current

The noise source can also be modeled by a current source in parallel with the resistor by taking the Norton equivalent that corresponds simply to divide by R. This gives the root mean square value of the current source as:

$$i_n = \sqrt{\frac{4k_B T \Delta f}{R}}.$$

Thermal noise is intrinsic to all resistors and is not a sign of poor design or manufacture, although resistors may also have excess noise.

Noise power in decibels

Signal power is often measured in dBm (decibels relative to 1 milliwatt). From the equation above, noise power in a resistor at room temperature, in dBm, is then:

$$P_{\mathrm{dBm}} = 10\ \log_{10}(k_B T \Delta f \times 1000)$$

where the factor of 1000 is present because the power is given in milliwatts, rather than watts. This equation can be simplified by separating the constant parts from the bandwidth:

$$P_{\mathrm{dBm}} = 10\ \log_{10}(k_B T \times 1000) + 10\ \log(\Delta f)$$

which is more commonly seen approximated as:

$$P_{\mathrm{dBm}} = -174 + 10\ \log_{10}(\Delta f)$$

Noise power at different bandwidths is then simple to calculate:

Bandwidth (Δf)	Thermal noise power	Notes
1 Hz	−174 dBm	
10 Hz	−164 dBm	
100 Hz	−154 dBm	
1 kHz	−144 dBm	
10 kHz	−134 dBm	FM channel of 2-way radio
100 kHz	−124 dBm	
180 kHz	−121.45 dBm	One LTE resource block
200 kHz	−120.98 dBm	One GSM channel (ARFCN)
1 MHz	−114 dBm	
2 MHz	−111 dBm	Commercial GPS channel
6 MHz	−106 dBm	Analog television channel
20 MHz	−101 dBm	WLAN 802.11 channel

Thermal noise on capacitors

Thermal noise on capacitors is referred to as *kTC* noise. Thermal noise in an RC circuit has an unusually simple expression, as the value of the resistance (*R*) drops out of the equation. This is because higher *R* contributes to more filtering as well as to more noise. The noise bandwidth of the RC circuit is 1/(4*RC*), which can substituted into the above formula to eliminate *R*. The mean-square and RMS noise voltage generated in such a filter are:

$$\bar{v}_n^2 = k_B T / C$$

$$v_n = \sqrt{k_B T/C}.$$

Thermal noise accounts for 100% of *kTC* noise, whether it is attributed to the resistance or to the capacitance.

In the extreme case of the *reset noise* left on a capacitor by opening an ideal switch, the resistance is infinite, yet the formula still applies; however, now the RMS must be interpreted not as a time average, but as an average over many such reset events, since the voltage is constant when the bandwidth is zero. In this sense, the Johnson noise of an RC circuit can be seen to be inherent, an effect of the thermodynamic distribution of the number of electrons on the capacitor, even without the involvement of a resistor.

The noise is not caused by the capacitor itself, but by the thermodynamic equilibrium of the amount of charge on the capacitor. Once the capacitor is disconnected from a conducting circuit, the thermodynamic fluctuation is *frozen* at a random value with standard deviation as given above.

The reset noise of capacitive sensors is often a limiting noise source, for example in image sensors. As an alternative to the voltage noise, the reset noise on the capacitor can also be quantified as the electrical charge standard deviation, as

$$Q_n = \sqrt{k_B T C}.$$

Since the charge variance is $k_B TC$, this noise is often called *kTC noise*.

Any system in thermal equilibrium has state variables with a mean energy of $kT/2$ per degree of freedom. Using the formula for energy on a capacitor ($E = ½CV^2$), mean noise energy on a capacitor can be seen to also be $½C(kT/C)$, or also $kT/2$. Thermal noise on a capacitor can be derived from this relationship, without consideration of resistance.

The *kTC* noise is the dominant noise source at small capacitors.

Noise of capacitors at 300 K

Capacitance	$\sqrt{k_B T/C}$	**Electrons**
1 fF	2 mV	12.5 e⁻
10 fF	640 µV	40 e⁻
100 fF	200 µV	125 e⁻
1 pF	64 µV	400 e⁻
10 pF	20 µV	1250 e⁻
100 pF	6.4 µV	4000 e⁻
1 nF	2 µV	12500 e⁻

Noise at very high frequencies

The above equations are good approximations at any practical radio frequency in use (i.e. frequencies below about 80 gigahertz). In the most general case, which includes up to optical frequencies, the power spectral density of the voltage across the resistor *R*, in V^2/Hz is given by:

$$\Phi(f) = \frac{2Rhf}{e^{\frac{hf}{k_B T}} - 1}$$

where f is the frequency, h Planck's constant, k_B Boltzmann constant and T the temperature in kelvins. If the frequency is low enough, that means:

$$f \ll \frac{k_B T}{h}$$

(this assumption is valid until few terahertz at room temperature) then the exponential can be expressed in terms of its Taylor series. The relationship then becomes:

$$\Phi(f) \approx 2Rk_B T.$$

In general, both *R* and *T* depend on frequency. In order to know the total noise it is enough to integrate over all the bandwidth. Since the signal is real, it is possible to integrate over only the positive frequencies, then multiply by 2. Assuming that *R* and *T* are constants over all the bandwidth Δf, then the root mean square (RMS) value of the voltage across a resistor due to thermal noise is given by

$$v_n = \sqrt{4k_B T R \Delta f},$$

that is, the same formula as above.

Chapter-10

Coherence

In physics, **coherence** is a property of waves that enables stationary (i.e. temporally and spatially constant) interference. More generally, coherence describes all properties of the correlation between physical quantities of a wave.

When interfering, two waves can add together to create a larger wave (**constructive interference**) or subtract from each other to create a smaller wave (**destructive interference**), depending on their relative phase. Two waves are said to be coherent if they have a constant relative phase. The degree of coherence is measured by the interference visibility, a measure of how perfectly the waves can cancel due to destructive interference.

Introduction

Coherence was originally introduced in connection with Young's double-slit experiment in optics but is now used in any field that involves waves, such as acoustics, electrical engineering, neuroscience, and quantum mechanics. The property of coherence is the basis for commercial applications such as holography, the Sagnac gyroscope, radio antenna arrays, optical coherence tomography and telescope interferometers (astronomical optical interferometers and radio telescopes).

Coherence and correlation

The coherence of two waves follows from how well correlated the waves are as quantified by the cross-correlation function. The cross-correlation quantifies the ability to predict the value of the second wave by knowing the value of the first. As an example, consider two waves perfectly correlated for all times. At any time, if the first wave changes, the second will change in the same way. If combined they can exhibit complete constructive interference/superposition at all times. It follows that they are perfectly coherent. As will be discussed below, the second wave need not be a separate entity. It could be the first wave at a different time or position. In this case, the measure of correlation is the autocorrelation function (sometimes called **self-coherence**). Degree of correlation involves correlation functions

Examples of wave-like states

These states are unified by the fact that their behavior is described by a wave equation or some generalization thereof.

- Waves in a rope (up and down) or slinky (compression and expansion)
- Surface waves in a liquid
- Electric signals (fields) in transmission cables
- Sound
- Radio waves and Microwaves
- Light waves (optics)
- Electrons, atoms and any other object (such as a baseball, as described by quantum physics)

In most of these systems, one can measure the wave directly. Consequently, its correlation with another wave can simply be calculated. However, in optics one cannot measure the electric field directly as it oscillates much faster than any detector's time resolution. Instead, we measure the intensity of the light. Most of the concepts involving coherence which will be introduced below were developed in the field of optics and then used in other fields. Therefore, many of the standard measurements of coherence are indirect measurements, even in fields where the wave can be measured directly.

Temporal coherence

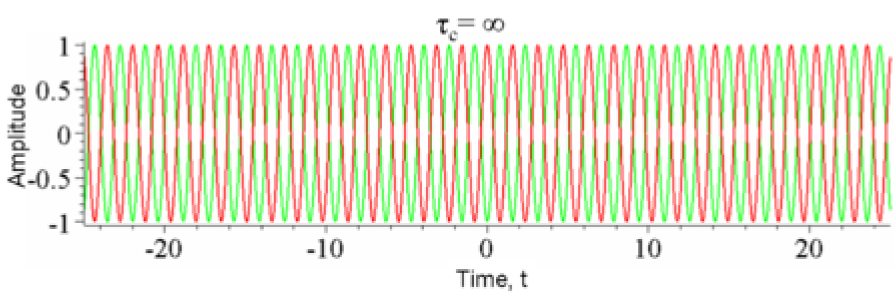


Figure 1: The amplitude of a single frequency wave as a function of time *t* (red) and a copy of the same wave delayed by τ(green). The coherence time of the wave is infinite since it is perfectly correlated with itself for all delays τ.

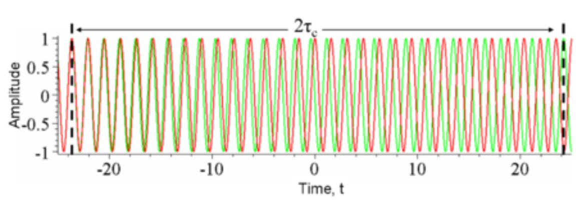


Figure 2: The amplitude of a wave whose phase drifts significantly in time τ_c as a function of time *t* (red) and a copy of the same wave delayed by $2\tau_c$(green). At any particular time t the wave can interfere perfectly with its delayed copy. But, since half the

time the red and green waves are in phase and half the time out of phase, when averaged over t any interference disappears at this delay.

Temporal coherence is the measure of the average correlation between the value of a wave at any pair of times, separated by delay τ. Temporal coherence tells us how monochromatic a source is. In other words, it characterizes how well a wave can interfere with itself at a different time. The delay over which the phase or amplitude wanders by a significant amount (and hence the correlation decreases by significant amount) is defined as the coherence time τ_c. At τ=0 the degree of coherence is perfect whereas it drops significantly by delay τ_c. The coherence length L_c is defined as the distance the wave travels in time τ_c.

One should be careful not to confuse the coherence time with the time duration of the signal, nor the coherence length with the coherence area.

The relationship between coherence time and bandwidth

It can be shown that the faster a wave decorrelates (and hence the smaller τ_c is) the larger the range of frequencies Δf the wave contains. Thus there is a tradeoff:

$$\tau_c \Delta f \approx 1.$$

Formally, this follows from the convolution theorem in mathematics, which relates the Fourier transform of the power spectrum (the intensity of each frequency) to its autocorrelation.

Examples of temporal coherence

We consider four examples of temporal coherence.

- A wave containing only a single frequency (monochromatic) is perfectly correlated at all times according to the above relation.
- Conversely, a wave whose phase drifts quickly will have a short coherence time.
- Similarly, pulses (wave packets) of waves, which naturally have a broad range of frequencies, also have a short coherence time since the amplitude of the wave changes quickly.
- Finally, white light, which has a very broad range of frequencies, is a wave which varies quickly in both amplitude and phase. Since it consequently has a very short coherence time (just 10 periods or so), it is often called incoherent.

The most monochromatic sources are usually lasers; such high monochromaticity implies long coherence lengths (up to hundreds of meters). For example, a stabilized helium-neon laser can produce light with coherence lengths in excess of 5 m. Not all lasers are monochromatic, however (e.g. for a mode-locked Ti-sapphire laser, $\Delta\lambda \approx 2$ nm - 70 nm). LEDs are characterized by $\Delta\lambda \approx 50$ nm, and tungsten filament lights exhibit $\Delta\lambda \approx 600$ nm, so these sources have shorter coherence times than the most monochromatic lasers.

Holography requires light with a long coherence time. In contrast, Optical coherence tomography uses light with a short coherence time.

Measurement of temporal coherence

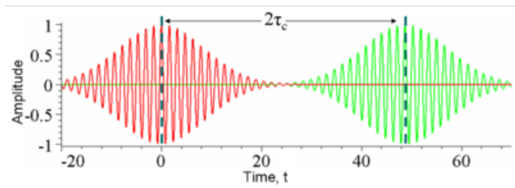


Figure 3: The amplitude of a wavepacket whose amplitude changes significantly in time τ_c (red) and a copy of the same wave delayed by $2\tau_c$(green) plotted as a function of time *t*. At any particular time the red and green waves are uncorrelated; one oscillates while the other is constant and so there will be no interference at this delay. Another way of looking at this is the wavepackets are not overlapped in time and so at any particular time there is only one nonzero field so no interference can occur.

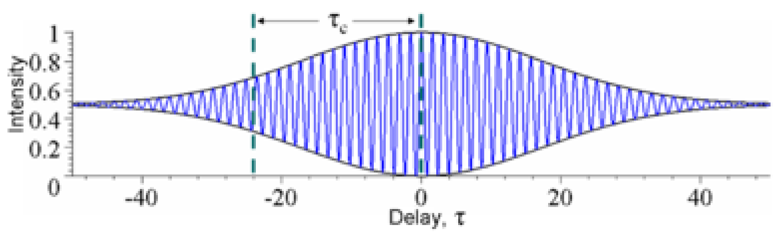


Figure 4: The time-averaged intensity (blue) detected at the output of an interferometer plotted as a function of delay τ for the example waves in Figures 2 and 3. As the delay is changed by half a period, the interference switches between constructive and destructive. The black lines indicate the interference envelope, which gives the degree of coherence. Although the waves in Figures 2 and 3 have different time durations, they have the same coherence time.

In optics, temporal coherence is measured in an interferometer such as the Michelson interferometer or Mach-Zehnder interferometer. In these devices, a wave is combined with a copy of itself that is delayed by time τ. A detector measures the time-averaged intensity of the light exiting the interferometer. The resulting interference visibility gives the temporal coherence at delay τ. Since for most natural light sources, the coherence time is much shorter than the time resolution of any detector, the detector itself does the time averaging. Consider the example shown in Figure 3. At a fixed delay, here $2\tau_c$, an infinitely fast detector would measure an intensity that fluctuates significantly over a time *t* equal to τ_c. In this case, to find the temporal coherence at $2\tau_c$, one would manually time-average the intensity.

Spatial coherence

In some systems, such as water waves or optics, wave-like states can extend over one or two dimensions. Spatial coherence describes the ability for two points in space, x_1 and x_2, in the extent of a wave to interfere, when averaged over time. More precisely, the spatial coherence is the cross-correlation between two points in a wave for all times. If a wave has only 1 value of amplitude over an infinite length, it is perfectly spatially coherent. The range of separation between the two points over which there is significant interference is called the coherence area, A_c. This is the relevant type of coherence for the Young's double-slit interferometer. It is also used in optical imaging systems and particularly in various types of astronomy telescopes. Sometimes people also use "spatial coherence" to refer to the visibility when a wave-like state is combined with a spatially shifted copy of itself.

Examples of spatial coherence

- Spatial coherence

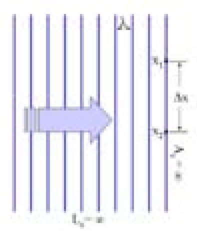


Figure 5: A plane wave with an infinite coherence length.

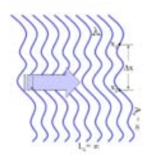


Figure 6: A wave with a varying profile (wavefront) and infinite coherence length.

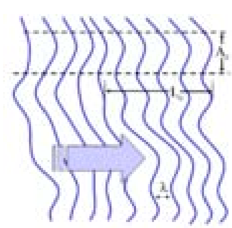


Figure 7: A wave with a varying profile (wavefront) and finite coherence length.

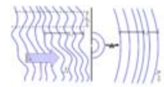


Figure 8: A wave with finite coherence area is incident on a pinhole (small aperture). The wave will diffract out of the pinhole. Far from the pinhole the emerging spherical wavefronts are approximately flat. The coherence area is now infinite while the coherence length is unchanged.

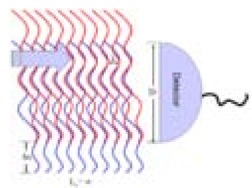


Figure 9: A wave with infinite coherence area is combined with a spatially-shifted copy of itself. Some sections in the wave interfere constructively and some will interfere destructively. Averaging over these sections, a detector with length D will measure reduced interference visibility. For example a misaligned Mach-Zehnder interferometer will do this.

Consider a tungsten light-bulb filament. Different points in the filament emit light independently and have no fixed phase-relationship. In detail, at any point in time the profile of the emitted light is going to be distorted. The profile will change randomly over the coherence time τ_c. Since for a white-light source such as a light-bulb τ_c is small, the filament is considered a spatially incoherent source. In contrast, a radio antenna array, has large spatial coherence because antennas at opposite ends of the array emit with a fixed phase-relationship. Light waves produced by a laser often have high temporal and spatial coherence (though the degree of coherence depends strongly on the exact properties of the laser). Spatial coherence of laser beams also manifests itself as speckle patterns and diffraction fringes seen at the edges of shadow.

Holography requires temporally and spatially coherent light. Its inventor, Dennis Gabor, produced successful holograms more than ten years before lasers were invented. To produce coherent light he passed the monochromatic light from an emission line of a mercury-vapor lamp through a pinhole spatial filter.

In February 2011, Dr Andrew Truscott, leader of a research team at the ARC Centre of Excellence for Quantum-Atom Optics at Australian National University in Canberra, Australian Capital Territory, showed that helium atoms cooled to near absolute zero / Bose-Einstein condensate state, can be made to flow and behave as a coherent beam as occurs in a laser.

Spectral coherence

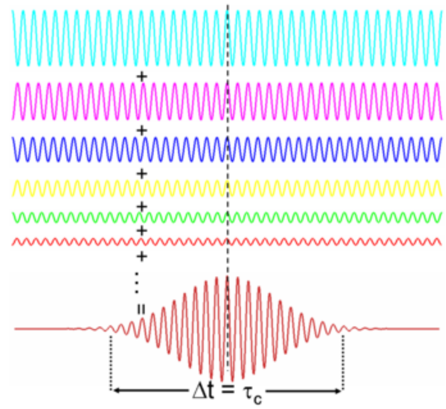


Figure 10: Waves of different frequencies (i.e. colors) interfere to form a pulse if they are coherent.

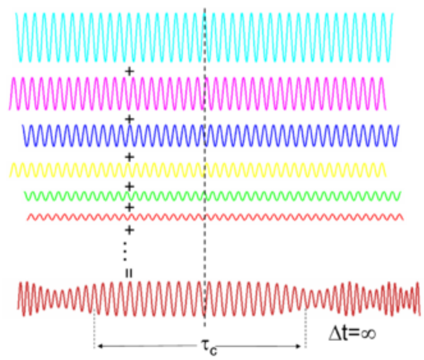


Figure 11: Spectrally incoherent light interferes to form continuous light with a randomly varying phase and amplitude

Waves of different frequencies (in light these are different colours) can interfere to form a pulse if they have a fixed relative phase-relationship. Conversely, if waves of different frequencies are not coherent, then, when combined, they create a wave that is continuous in time (e.g. white light or white noise). The temporal duration of the pulse Δt is limited by the spectral bandwidth of the light Δf according to:

$$\Delta f \Delta t \geq 1,$$

which follows from the properties of the Fourier transform (for quantum particles it also results in the Heisenberg uncertainty principle).

If the phase depends linearly on the frequency (i.e. $\theta(f) \propto f$) then the pulse will have the minimum time duration for its bandwidth (a *transform-limited* pulse), otherwise it is chirped.

Measurement of spectral coherence

Measurement of the spectral coherence of light requires a nonlinear optical interferometer, such as an intensity optical correlator, frequency-resolved optical gating (FROG), or Spectral phase interferometry for direct electric-field reconstruction (SPIDER).

Polarization coherence

Light also has a polarization, which is the direction in which the electric field oscillates. Unpolarized light is composed of two equally intense incoherent light waves with orthogonal polarizations. The electric field of the unpolarized light wanders in every direction and changes in phase over the coherence time of the two light waves. A polarizer rotated to any angle will always transmit half the incident intensity when averaged over time.

If the electric field wanders by a smaller amount the light will be partially polarized so that at some angle, the polarizer will transmit more than half the intensity. If a wave is combined with an orthogonally polarized copy of itself delayed by less than the coherence time, partially polarized light is created.

The polarization of a light beam is represented by a vector in the Poincare sphere. For polarized light the end of the vector lies on the surface of the sphere, whereas the vector has zero length for unpolarized light. The vector for partially polarized light lies within the sphere

Applications

Holography

Coherent superpositions of *optical wave fields* include holography. Holographic objects are used frequently in daily life in bank notes and credit cards.

Non-optical wave fields

Further applications concern the coherent superposition of *non-optical wave fields*. In quantum mechanics for example one considers a probability field, which is related to the wave function $\psi(\mathbf{r})$(interpretation: density of the probability amplitude). Here the applications concern, among others, the future technologies of quantum computing and the already available technology of quantum cryptography. Additionally the problems of the following subchapter are treated.

Quantum coherence

In quantum mechanics, all objects have wave-like properties. For instance, in Young's double-slit experiment electrons can be used in the place of light waves. Each electron can go through either slit and hence has two paths that it can take to a particular final position. In quantum mechanics these two paths interfere. If there is destructive interference, the electron never arrives at that particular position. This ability to interfere is called quantum coherence.

The quantum description of perfectly coherent paths is called a pure state, in which the two paths are combined in a superposition. The correlation between the two particles exceeds what would be predicted for classical correlation alone. If this two-particle system is decohered (which would occur in a measurement via Einselection), then there is no longer any phase relationship between the two states. The quantum description of imperfectly coherent paths is called a mixed state, described by a density matrix (also called the "statistical operator") and is entirely analogous to a classical system of mixed probabilities (the correlations are classical).

Large-scale (macroscopic) quantum coherence leads to novel phenomena. For instance, the laser, superconductivity, and superfluidity are examples of highly coherent quantum systems, whose effects are evident at the macroscopic scale. These examples of quantum coherence are Bose–Einstein condensates. Here, all the particles that make up the condensate are in-phase; they are thus necessarily all described by a single quantum wavefunction. On the other hand, the Schrödinger's cat thought experiment, highlights the fact that quantum coherence is not typically seen at the macroscopic scale but has been observed in the motion of mechanical resonator.

Chapter-11

Sonar

French F70 type frigates (here, *La Motte-Picquet*) are fitted with VDS (Variable Depth Sonar) type DUBV43 or DUBV43C towed sonars

Sonar (originally an acronym for **SO**und **N**avigation **A**nd **R**anging) is a technique that uses sound propagation (usually underwater, as in Submarine navigation) to navigate, communicate with or detect other vessels. Two types of technology share the name "sonar": *passive* sonar is essentially listening for the sound made by vessels; *active* sonar is emitting pulses of sounds and listening for echoes. Sonar may be used as a means of acoustic location and of measurement of the echo characteristics of "targets" in the water. Acoustic location in air was used before the introduction of radar. Sonar may also be used in air for robot navigation, and SODAR (an upward looking in-air sonar) is used for atmospheric investigations. The term *sonar* is also used for the equipment used to generate and receive the sound. The acoustic frequencies used in sonar systems vary from very low (infrasonic) to extremely high (ultrasonic). The study of underwater sound is known as underwater acoustics or hydroacoustics.

History

Although some animals (dolphins and bats) have used sound for communication and object detection for millions of years, use by humans in the water is initially recorded by Leonardo Da Vinci in 1490: a tube inserted into the water was said to be used to detect vessels by placing an ear to the tube.

In the 19th century an underwater bell was used as an ancillary to lighthouses to provide warning of hazards.

The use of sound to 'echo locate' underwater in the same way as bats use sound for aerial navigation seems to have been prompted by the *Titanic* disaster of 1912. The world's first patent for an underwater echo ranging device was filed at the British Patent Office by English meteorologist Lewis Richardson a month after the sinking of the Titanic, and a German physicist Alexander Behm obtained a patent for an echo sounder in 1913.

The Canadian engineer Reginald Fessenden, while working for the Submarine Signal Company in Boston, built an experimental system beginning in 1912, a system later tested in Boston Harbor, and finally in 1914 from the U.S. Revenue (now Coast Guard) Cutter Miami on the Grand Banks off Newfoundland Canada. In that test, Fessenden demonstrated depth sounding, underwater communications (Morse Code) and echo ranging (detecting an iceberg at two miles (3 km) range). The so-called Fessenden oscillator, at ca. 500 Hz frequency, was unable to determine the bearing of the berg due to the 3 metre wavelength and the small dimension of the transducer's radiating face (less than 1 metre in diameter). The ten Montreal-built British H class submarines launched in 1915 were equipped with a Fessenden oscillator.

During World War I the need to detect submarines prompted more research into the use of sound. The British made early use of underwater hydrophones, while the French physicist Paul Langevin, working with a Russian immigrant electrical engineer, Constantin Chilowski, worked on the development of active sound devices for detecting submarines in 1915 using quartz. Although piezoelectric and magnetostrictive transducers later superseded the electrostatic transducers they used, this work influenced future designs. Lightweight sound-sensitive plastic film and fibre optics have been used for hydrophones (acousto-electric transducers for in-water use), while Terfenol-D and PMN (lead magnesium niobate) have been developed for projectors.

ASDIC

In 1916, under the British Board of Invention and Research, Canadian physicist Robert William Boyle took on the active sound detection project with A B Wood, producing a prototype for testing in mid 1917. This work, for the Anti-Submarine Division of the British Naval Staff, was undertaken in utmost secrecy, and used quartz piezoelectric crystals to produce the world's first practical underwater active sound detection apparatus. To maintain secrecy no mention of sound experimentation or quartz was made - the word used to describe the early work ('supersonics') was changed to 'ASD'ics, and the quartz material to 'ASD'ivite: hence the British acronym *ASDIC*. In 1939, in response to a question from the Oxford English Dictionary, the Admiralty made up the story that it stood for 'Allied Submarine Detection Investigation Committee', and this is still widely believed, though no committee bearing this name has been found in the Admiralty archives.

By 1918, both France and Britain had built prototype active systems. The British tested their ASDIC on HMS *Antrim* in 1920, and started production in 1922. The 6th Destroyer Flotilla had ASDIC-equipped vessels in 1923. An anti-submarine school, HMS *Osprey*, and a training flotilla of four vessels were established on Portland in 1924. The US Sonar QB set arrived in 1931.

By the outbreak of World War II, the Royal Navy had five sets for different surface ship classes, and others for submarines, incorporated into a complete anti-submarine attack system. The effectiveness of early ASDIC was hamstrung by the use of the depth charge as an anti-submarine weapon. This required an attacking vessel to pass over a submerged contact before dropping charges over the stern, resulting in a loss of ASDIC contact in the moments leading up to attack. The hunter was effectively firing blind, during which time a submarine commander could take evasive action. This situation was remedied by using several ships cooperating and by the adoption of "ahead throwing weapons", such as Hedgehog and later Squid, which projected warheads at a target ahead of the attacker and thus still in ASDIC contact. Developments during the war resulted in British ASDIC sets which used several different shapes of beam, continuously covering blind spots. Later, acoustic torpedoes were used.

At the start of World War II, British ASDIC technology was transferred for free to the United States. Research on ASDIC and underwater sound was expanded in the UK and in the US. Many new types of military sound detection were developed. These included sonobuoys, first developed by the British in 1944 under the codename *High Tea*, dipping/dunking sonar and mine detection sonar. This work formed the basis for post war developments related to countering the nuclear submarine. Work on sonar had also been carried out in the Axis countries, notably in Germany, which included countermeasures. At the end of World War II this German work was assimilated by Britain and the US. Sonars have continued to be developed by many countries, including Russia, for both military and civil uses. In recent years the major military development has been the increasing interest in low frequency active systems.

SONAR

During the 1930s American engineers developed their own underwater sound detection technology and important discoveries were made, such as thermoclines, that would help future development. After technical information was exchanged between the two countries during the Second World War, Americans began to use the term *SONAR* for their systems, coined as the equivalent of RADAR.

Performance factors

The detection, classification and localisation performance of a sonar depends on the environment and the receiving equipment, as well as the transmitting equipment in an active sonar or the target radiated noise in a passive sonar.

Sound propagation

Sonar operation is affected by variations in sound speed, particularly in the vertical plane. Sound travels more slowly in fresh water than in sea water, though the difference is small. The speed is determined by the water's bulk modulus and mass density. The bulk modulus is affected by temperature, dissolved impurities (usually salinity), and pressure. The density effect is small. The speed of sound (in feet per second) is approximately:

> 4388 + (11.25 × temperature (in °F)) + (0.0182 × depth (in feet)) + salinity (in parts-per-thousand).

This empirically derived approximation equation is reasonably accurate for normal temperatures, concentrations of salinity and the range of most ocean depths. Ocean temperature varies with depth, but at between 30 and 100 meters there is often a marked change, called the thermocline, dividing the warmer surface water from the cold, still waters that make up the rest of the ocean. This can frustrate sonar, because a sound originating on one side of the thermocline tends to be bent, or refracted, through the thermocline. The thermocline may be present in shallower coastal waters. However, wave action will often mix the water column and eliminate the thermocline. Water pressure also affects sound propagation: higher pressure increases the sound speed, which causes the sound waves to refract away from the area of higher sound speed. The mathematical model of refraction is called Snell's law.

If the sound source is deep and the conditions are right, propagation may occur in the 'deep sound channel'. This provides extremely low propagation loss to a receiver in the channel. This is because of sound trapping in the channel with no losses at the boundaries. Similar propagation can occur in the 'surface duct' under suitable conditions. However in this case there are reflection losses at the surface.

In shallow water propagation is generally by repeated reflection at the surface and bottom, where considerable losses can occur.

Sound propagation is affected by absorption in the water itself as well as at the surface and bottom. This absorption depends upon frequency, with several different mechanisms in sea water. Long-range sonar uses low frequencies to minimise absorption effects.

The sea contains many sources of noise that interfere with the desired target echo or signature. The main noise sources are waves and shipping. The motion of the receiver through the water can also cause speed-dependent low frequency noise.

Scattering

When active sonar is used, scattering occurs from small objects in the sea as well as from the bottom and surface. This can be a major source of interference. This acoustic scattering is analogous to the scattering of the light from a car's headlights in fog: a high-intensity pencil beam will penetrate the fog to some extent, but broader-beam headlights emit much light in unwanted directions, much of which is scattered back to the observer, overwhelming that reflected from the target ("white-out"). For analogous reasons active sonar needs to transmit in a narrow beam to minimise scattering.

Target characteristics

The sound *reflection* characteristics of the target of an active sonar, such as a submarine, are known as its target strength. A complication is that echoes are also obtained from other objects in the sea such as whales, wakes, schools of fish and rocks.

Passive sonar detects the target's *radiated* noise characteristics. The radiated spectrum comprises a continuous spectrum of noise with peaks at certain frequencies which can be used for classification.

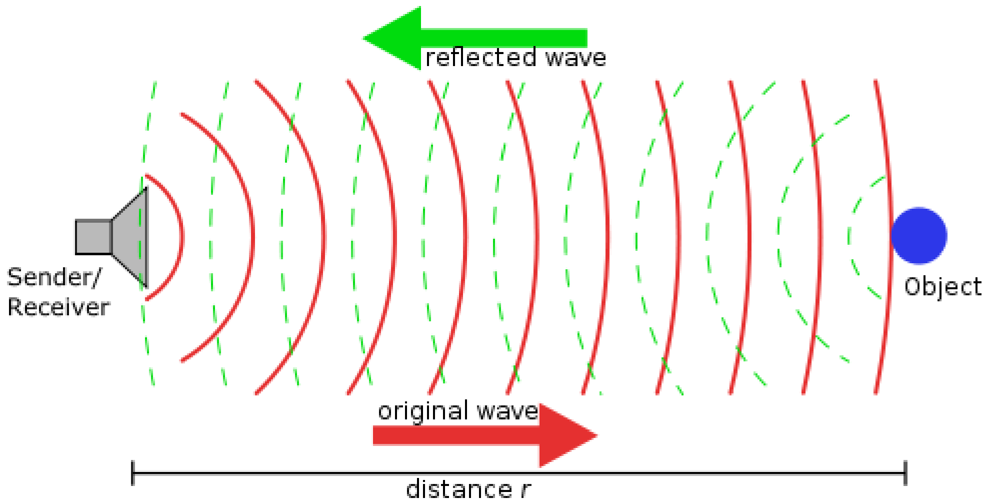
Countermeasures

Active (powered) countermeasures may be launched by a submarine under attack to raise the noise level, provide a large false target, and obscure the signature of the submarine itself.

Passive (i.e., non-powered) countermeasures include:

- Mounting noise-generating devices on isolating devices.
- Sound-absorbent coatings on the hulls of submarines, for example anechoic tiles.

Active sonar



Principle of an active sonar

Active sonar uses a sound transmitter and a receiver. When the two are in the same place it is monostatic operation. When the transmitter and receiver are separated it is bistatic operation. When more transmitters (or more receivers) are used, again spatially separated, it is multistatic operation. Most sonars are used monostatically with the same array often being used for transmission and reception. Active sonobuoy fields may be operated multistatically.

Active sonar creates a pulse of sound, often called a "ping", and then listens for reflections (echo) of the pulse. This pulse of sound is generally created electronically using a sonar Projector consisting of a signal generator, power amplifier and electro-acoustic transducer/array. A beamformer is usually employed to concentrate the acoustic power into a beam, which may be swept to cover the required search angles. Generally, the electro-acoustic transducers are of the Tonpilz type and their design may be optimised to achieve maximum efficiency over the widest bandwidth, in order to optimise performance of the overall system. Occasionally, the acoustic pulse may be created by other means, e.g. (1) chemically using explosives, or (2) airguns or (3) plasma sound sources.

To measure the distance to an object, the time from transmission of a pulse to reception is measured and converted into a range by knowing the speed of sound. To measure the bearing, several hydrophones are used, and the set measures the relative arrival time to each, or with an array of hydrophones, by measuring the relative amplitude in beams formed through a process called beamforming. Use of an array reduces the spatial response so that to provide wide cover multibeam systems are used. The target signal (if present) together with noise is then passed through various forms of signal processing, which for simple sonars may be just energy measurement. It is then presented to some form of decision device that calls the output either the required signal or noise. This decision device may be an operator with headphones or a display, or in more sophisticated sonars this function may be carried out by software. Further processes may be carried out to classify the target and localise it, as well as measuring its velocity.

The pulse may be at constant frequency or a chirp of changing frequency (to allow pulse compression on reception). Simple sonars generally use the former with a filter wide enough to cover possible Doppler changes due to target movement, while more complex ones generally include the latter technique. Since digital processing became available pulse compression has usually been implemented using digital correlation techniques. Military sonars often have multiple beams to provide all-round cover while simple ones only cover a narrow arc, although the beam may be rotated, relatively slowly, by mechanical scanning.

Particularly when single frequency transmissions are used, the Doppler effect can be used to measure the radial speed of a target. The difference in frequency between the transmitted and received signal is measured and converted into a velocity. Since Doppler shifts can be introduced by either receiver or target motion, allowance has to be made for the radial speed of the searching platform.

One useful small sonar is similar in appearance to a waterproof flashlight. The head is pointed into the water, a button is pressed, and the device displays the distance to the target. Another variant is a "fishfinder" that shows a small display with shoals of fish. Some civilian sonars (which are not designed for stealth) approach active military sonars in capability, with quite exotic three-dimensional displays of the area near the boat.

When active sonar is used to measure the distance from the transducer to the bottom, it is known as echo sounding. Similar methods may be used looking upward for wave measurement.

Active sonar is also used to measure distance through water between two sonar transducers or a combination of a hydrophone (underwater acoustic microphone) and projector (underwater acoustic speaker). A transducer is a device that can transmit and receive acoustic signals ("pings"). When a hydrophone/transducer receives a specific interrogation signal it responds by transmitting a specific reply signal. To measure distance, one transducer/projector transmits an interrogation signal and measures the time between this transmission and the receipt of the other transducer/hydrophone reply. The time difference, scaled by the speed of sound through water and divided by two, is the distance between the two platforms. This technique, when used with multiple transducers/hydrophones/projectors, can calculate the relative positions of static and moving objects in water.

In combat situations, an active pulse can be detected by an opponent and will reveal a submarine's position.

A very directional, but low-efficiency, type of sonar (used by fisheries, military, and for port security) makes use of a complex nonlinear feature of water known as non-linear sonar, the virtual transducer being known as a *parametric array*.

Project ARTEMIS

Project ARTEMIS was a one-of-a-kind low-frequency sonar for surveillance that was deployed off Bermuda for several years in the early 1960s. The active portion was deployed from a World War II tanker, and the receiving array was a built into a fixed position on an offshore bank.

Transponder

This is an active sonar device that receives a stimulus and immediately (or with a delay) retransmits the received signal or a predetermined one.

Performance prediction

A sonar target is small relative to the sphere, centred around the emitter, on which it is located. Therefore, the power of the reflected signal is very low, several orders of magnitude less than the original signal. Even if the reflected signal was of the same power, the following example (using hypothetical values) shows the problem: Suppose a sonar system is capable of emitting a 10,000 W/m² signal at 1 m, and detecting a 0.001 W/m² signal. At 100 m the signal will be 1 W/m² (due to the inverse-square law). If the entire signal is reflected from a 10 m² target, it will be at 0.001 W/m² when it reaches the emitter, i.e. just detectable. However, the original signal will remain above 0.001 W/m² until 300 m. Any 10 m² target between 100 and 300 m using a similar or better system would be able to detect the pulse but would not be detected by the emitter. The detectors must be very sensitive to pick up the echoes. Since the original signal is much more powerful, it can be detected many times further than twice the range of the sonar (as in the example).

In active sonar there are two performance limitations, due to noise and reverberation. In general one or other of these will dominate so that the two effects can be initially considered separately.

In noise limited conditions at initial detection:

$$\mathrm{SL} - 2\mathrm{TL} + \mathrm{TS} - (\mathrm{NL} - \mathrm{DI}) = \mathrm{DT}$$

where SL is the source level, TL is the transmission loss (or propagation loss), TS is the target strength, NL is the noise level, DI is the directivity index of the array (an approximation to the array gain) and DT is the detection threshold.

In reverberation limited conditions at initial detection (neglecting array gain):

$$\mathrm{SL} - 2\mathrm{TL} + \mathrm{TS} = \mathrm{RL} + \mathrm{DT}$$

where RL is the reverberation level and the other factors are as before.

Marine mammals

A Humpback whale

Active sonar may harm marine animals, although the precise mechanisms for this are not well understood. Some marine animals, such as whales and dolphins, use echolocation systems, sometimes called *biosonar* to locate predators and prey. It is conjectured that active sonar transmitters could confuse these animals and interfere with basic biological functions such as feeding and mating.

Hand-held sonar for use by a diver

Scuba diver using INSS hand-held sonar

- The LIMIS (= Limpet Mine Imaging Sonar) is a hand-held or ROV-mounted imaging sonar for use by a diver. Its name is because it was designed for patrol divers (combat frogmen or Clearance Divers) to look for limpet mines in low visibility water. Links:

    - Abstract of article by the International Society for Optical Engineering
    - Used to find debris from the Space Shuttle Columbia crash
    - Used in fish passage research at hydropower facilities
- The LUIS (= Lensing Underwater Imaging System) is another imaging sonar for use by a diver. Links:
    - Used for counting salmon in a river
- There is or was a small flashlight-shaped handheld sonar for divers, that merely displays range.
- For the INSS = Integrated Navigation Sonar System see:
    - an image.
    - short description
    - description

Passive sonar

Passive sonar listens without transmitting. It is often employed in military settings, although it is also used in science applications, *e.g.*, detecting fish for presence/absence studies in various aquatic environments. In the very broadest usage, this term can encompass virtually any analytical technique involving remotely generated sound, though it is usually restricted to techniques applied in an aquatic environment.

Identifying sound sources

Passive sonar has a wide variety of techniques for identifying the source of a detected sound. For example, U.S. vessels usually operate 60 Hz alternating current power systems. If transformers or generators are mounted without proper vibration insulation from the hull or become flooded, the 60 Hz sound from the windings can be emitted from the submarine or ship. This can help to identify its nationality, as most European submarines have 50 Hz power systems. Intermittent sound sources (such as a wrench being dropped) may also be detectable to passive sonar. Until fairly recently, an experienced trained operator identified signals, but now computers may do this.

Passive sonar systems may have large sonic databases, but the sonar operator usually finally classifies the signals manually. A computer system frequently uses these databases to identify classes of ships, actions (i.e. the speed of a ship, or the type of weapon released), and even particular ships. Publications for classification of sounds are provided by and continually updated by the US Office of Naval Intelligence.

Noise limitations

Passive sonar on vehicles is usually severely limited because of noise generated by the vehicle. For this reason, many submarines operate nuclear reactors that can be cooled without pumps, using silent convection, or fuel cells or batteries, which can also run silently. Vehicles' propellers are also designed and precisely machined to emit minimal noise. High-speed propellers often create tiny bubbles in the water, and this cavitation has a distinct sound.

The sonar hydrophones may be towed behind the ship or submarine in order to reduce the effect of noise generated by the watercraft itself. Towed units also combat the thermocline, as the unit may be towed above or below the thermocline.

The display of most passive sonars used to be a two-dimensional waterfall display. The horizontal direction of the display is bearing. The vertical is frequency, or sometimes time. Another display technique is to color-code frequency-time information for bearing. More recent displays are generated by the computers, and mimic radar-type plan position indicator displays.

Performance prediction

Unlike active sonar, only one way propagation is involved. Because of the different signal processing used, the minimum detectable signal to noise ratio will be different. The equation for determining the performance of a passive sonar is:

$$SL - TL = NL - DI + DT$$

where SL is the source level, TL is the transmission loss, NL is the noise level, DI is the directivity index of the array (an approximation to the array gain) and DT is the detection threshold. The figure of merit of a passive sonar is:

$$FOM = SL + DI - (NL + DT).$$

Warfare

Modern naval warfare makes extensive use of both passive and active sonar from water-borne vessels, aircraft and fixed installations. The relative usefulness of active versus passive sonar depends on the radiated noise characteristics of the target, generally a submarine. Although in World War II active sonar was used by surface craft—submarines avoided emitting pings which revealed their presence and position—with the advent of modern signal-processing passive sonar became preferred for initial detection. Submarines were then designed for quieter operation, and active sonar is now more used. In 1987 a division of Japanese company Toshiba reportedly sold machinery to the Soviet Union that allowed it to mill submarine propeller blades so that they became radically quieter, creating a huge security issue with their newer generation of submarines.

Active sonar gives the exact bearing to a target, and sometimes the range. Active sonar works the same way as radar: a signal is emitted. The sound wave then travels in many directions from the emitting object. When it hits an object, the sound wave is then reflected in many other directions. Some of the energy will travel back to the emitting source. The echo will enable the sonar system or technician to calculate, with many factors such as the frequency, the energy of the received signal, the depth, the water temperature, the position of the reflecting object, etc. Active sonar is used when the platform commander determines that it is more important to determine the position of a possible threat submarine than it is to conceal his own position. With surface ships it might be assumed that the threat is already tracking the ship with satellite data. Any vessel around the emitting sonar will detect the emission. Having heard the signal, it is easy to identify the sonar equipment used (usually with its frequency) and its position (with the sound wave's energy). Active sonar is similar to radar in that, while it allows detection of targets at a certain range, it also enables the emitter to be detected at a far greater range, which is undesirable.

Since active sonar reveals the presence and position of the operator, and does not allow exact classification of targets, it is used by fast (planes, helicopters) and by noisy platforms (most surface ships) but rarely by submarines. When active sonar is used by surface ships or submarines, it is typically activated very briefly at intermittent periods to minimise the risk of detection. Consequently active sonar is normally considered a backup to passive sonar. In aircraft, active sonar is used in the form of disposable sonobuoys that are dropped in the aircraft's patrol area or in the vicinity of possible enemy sonar contacts.

Passive sonar has several advantages. Most importantly, it is silent. If the target radiated noise level is high enough, it can have a greater range than active sonar, and allows the target to be identified. Since any motorized object makes some noise, it may in principle be detected, depending on the level of noise emitted and the ambient noise level in the area, as well as the technology used. To simplify, passive sonar "sees" around the ship using it. On a submarine, nose-mounted passive sonar detects in directions of about 270°, centered on the ship's alignment, the hull-mounted array of about 160° on each side, and the towed array of a full 360°. The invisible areas are due to the ship's own interference. Once a signal is detected in a certain direction (which means that something makes sound in that direction, this is called broadband detection) it is possible to zoom in and analyze the signal received (narrowband analysis). This is generally done using a Fourier transform to show the different frequencies making up the sound. Since every engine makes a specific sound, it is straightforward to identify the object. Databases of unique engine sounds are part of what is known as *acoustic intelligence* or ACINT.

Another use of passive sonar is to determine the target's trajectory. This process is called Target Motion Analysis (TMA), and the resultant "solution" is the target's range, course, and speed. TMA is done by marking from which direction the sound comes at different times, and comparing the motion with that of the operator's own ship. Changes in relative motion are analyzed using standard geometrical techniques along with some assumptions about limiting cases.

Passive sonar is stealthy and very useful. However, it requires high-tech electronic components and is costly. It is generally deployed on expensive ships in the form of arrays to enhance detection. Surface ships use it to good effect; it is even better used by submarines, and it is also used by airplanes and helicopters, mostly to a "surprise effect", since submarines can hide under thermal layers. If a submarine's commander believes he is alone, he may bring his boat closer to the surface and be easier to detect, or go deeper and faster, and thus make more sound.

Examples of sonar applications in military use are given below. Many of the civil uses given in the following section may also be applicable to naval use.

Anti-submarine warfare

Variable Depth Sonar and its winch

Until recently, ship sonars were usually with hull mounted arrays, either amidships or at the bow. It was soon found after their initial use that a means of reducing flow noise was required. The first were made of canvas on a framework, then steel ones were used. Now domes are usually made of reinforced plastic or pressurised rubber. Such sonars are primarily active in operation. An example of a conventional hull mounted sonar is the SQS-56.

Because of the problems of ship noise, towed sonars are also used. These also have the advantage of being able to be placed deeper in the water. However, there are limitations on their use in shallow water. These are called towed arrays (linear) or variable depth sonars (VDS) with 2/3D arrays. A problem is that the winches required to deploy/recover these are large and expensive. VDS sets are primarily active in operation while towed arrays are passive.

An example of a modern active/passive ship towed sonar is Sonar 2087 made by Thales Underwater Systems.

Torpedoes

Modern torpedoes are generally fitted with an active/passive sonar. This may be used to home directly on the target, but wake following torpedoes are also used. An early example of an acoustic homer was the Mark 37 torpedo.

Torpedo countermeasures can be towed or free. An early example was the German Sieglinde device while the Pillenwerfer was a chemical device. A widely used US device was the towed Nixie while MOSS submarine simulator was a free device. A modern alternative to the Nixie system is the UK Royal Navy S2170 Surface Ship Torpedo Defence system.

Mines

Mines may be fitted with a sonar to detect, localize and recognize the required target.

Mine countermeasures

Mine Countermeasure (MCM) Sonar, sometimes called "Mine and Obstacle Avoidance Sonar (MOAS)", is a specialised type of sonar used for detecting small objects. Most MCM sonars are hull mounted but a few types are VDS design. An example of a hull mounted MCM sonar is the Type 2193 while the SQQ-32 Mine-hunting sonar and Type 2093 systems are VDS designs.

Submarine navigation

Submarines rely on sonar to a greater extent than surface ships as they cannot use radar at depth. The sonar arrays may be hull mounted or towed. Information fitted on typical fits is given in Oyashio class submarine and Swiftsure class submarine.

Aircraft

Helicopters can be used for antisubmarine warfare by deploying fields of active/passive sonobuoys or can operate dipping sonar, such as the AQS-13. Fixed wing aircraft can also deploy sonobuoys and have greater endurance and capacity to deploy them. Processing from the sonobuoys or dipping sonar can be on the aircraft or on ship.

Helicopters have also been used for mine countermeasure missions using towed sonars such as the AQS-20A

AN/AQS-13 Dipping sonar deployed from an H-3 Sea King.

Underwater communications

Dedicated sonars can be fitted to ships and submarines for underwater communication.

Ocean surveillance

For many years, the United States operated a large set of passive sonar arrays at various points in the world's oceans, collectively called Sound Surveillance System (SOSUS) and later Integrated Undersea Surveillance System (IUSS). A similar system is believed to have been operated by the Soviet Union. As permanently mounted arrays in the deep ocean were utilised, they were in very quiet conditions so long ranges could be achieved. Signal processing was carried out using powerful computers ashore. With the ending of the Cold War a SOSUS array has been turned over to scientific use.

In the United States Navy, a special badge known as the Integrated Undersea Surveillance System Badge is awarded to those who have been trained and qualified in its operation.

Underwater security

Sonar can be used to detect frogmen and other scuba divers. This can be applicable around ships or at entrances to ports. Active sonar can also be used as a deterrent and/or disablement mechanism. One such device is the Cerberus system.

Hand-held sonar

Limpet Mine Imaging Sonar (LIMIS) is a hand-held or ROV-mounted imaging sonar designed for patrol divers (combat frogmen or clearance divers) to look for limpet mines in low visibility water.

The LUIS is another imaging sonar for use by a diver.

Integrated Navigation Sonar System (INSS) is a small flashlight-shaped handheld sonar for divers that displays range.

Intercept sonar

This is a sonar designed to detect and locate the transmissions from hostile active sonars. An example of this is the Type 2082 fitted on the British Vanguard class submarines.

Civilian applications

Fisheries

Fishing is an important industry that is seeing growing demand, but world catch tonnage is falling as a result of serious resource problems. The industry faces a future of continuing worldwide consolidation until a point of sustainability can be reached. However, the consolidation of the fishing fleets are driving increased demands for sophisticated fish finding electronics such as sensors, sounders and sonars. Historically, fishermen have used many different techniques to find and harvest fish. However, acoustic technology has been one of the most important driving forces behind the development of the modern commercial fisheries.

Sound waves travel differently through fish than through water because a fish's air-filled swim bladder has a different density than seawater. This density difference allows the detection of schools of fish by using reflected sound. Acoustic technology is especially well suited for underwater applications since sound travels farther and faster underwater than in air. Today, commercial fishing vessels rely almost completely on acoustic sonar and sounders to detect fish. Fishermen also use active sonar and echo sounder technology to determine water depth, bottom contour, and bottom composition.

Cabin display of a fish finder sonar

Companies such as Raymarine UK, Marport Canada, Wesmar, Furuno, Krupp, and Simrad make a variety of sonar and acoustic instruments for the deep sea commercial fishing industry. For example, net sensors take various underwater measurements and transmit the information back to a receiver onboard a vessel. Each sensor is equipped with one or more acoustic transducers depending on its specific function. Data is transmitted from the sensors using wireless acoustic telemetry and is received by a hull mounted hydrophone. The analog signals are decoded and converted by a digital acoustic receiver into data which is transmitted to a bridge computer for graphical display on a high resolution monitor.

Echo sounding

An echo-sounder sends an acoustic pulse directly downwards to the seabed and records the returned echo. The sound pulse is generated by a transducer that emits an acoustic pulse and then “listens” for the return signal. The time for the signal to return is recorded and converted to a depth measurement by calculating the speed of sound in water. As the speed of sound in water is around 1,500 metres per second, the time interval, measured in milliseconds, between the pulse being transmitted and the echo being received, allows bottom depth and targets to be measured.

The value of underwater acoustics to the fishing industry has led to the development of other acoustic instruments that operate in a similar fashion to echo-sounders but, because their function is slightly different from the initial model of the echo-sounder, have been given different terms.

Net location

The net sounder is an echo sounder with a transducer mounted on the headline of the net rather than on the bottom of the vessel. Nevertheless, to accommodate the distance from the transducer to the display unit, which is much greater than in a normal echo-sounder, several refinements have to be made. Two main types are available. The first is the cable type in which the signals are sent along a cable. In this case there has to be the provision of a cable drum on which to haul, shoot and stow the cable during the different phases of the operation. The second type is the cable less net-sounder – such as Marport's Trawl Explorer - in which the signals are sent acoustically between the net and hull mounted receiver/hydrophone on the vessel. In this case no cable drum is required but sophisticated electronics are needed at the transducer and receiver.

The display on a net sounder shows the distance of the net from the bottom (or the surface), rather than the depth of water as with the echo-sounder's hull-mounted transducer. Fixed to the headline of the net, the footrope can usually be seen which gives an indication of the net performance. Any fish passing into the net can also be seen, allowing fine adjustments to be made to catch the most fish possible. In other fisheries, where the amount of fish in the net is important, catch sensor transducers are mounted at various positions on the cod-end of the net. As the cod-end fills up these catch sensor transducers are triggered one by one and this information is transmitted acoustically to display monitors on the bridge of the vessel. The skipper can then decide when to haul the net.

Modern versions of the net sounder, using multiple element transducers, function more like a sonar than an echo sounder and show slices of the area in front of the net and not merely the vertical view that the initial net sounders used.

The sonar is an echo-sounder with a directional capability that can show fish or other objects around the vessel.good

Ship velocity measurement

Sonars have been developed for measuring a ship's velocity either relative to the water or to the bottom.

ROV and UUV

Small sonars have been fitted to Remotely Operated Vehicles (ROV) and Unmanned Underwater Vehicles (UUV) to allow their operation in murky conditions. These sonars are used for looking ahead of the vehicle. The Long-Term Mine Reconnaissance System is an UUV for MCM purposes.

Vehicle location

Sonars which act as beacons are fitted to aircraft to allow their location in the event of a crash in the sea. Short and Long Baseline sonars may be used for caring out the location, such as LBL.

Scientific applications

Biomass estimation

Detection of fish, and other marine and aquatic life, and estimation their individual sizes or total biomass using active sonar techniques. As the sound pulse travels through water it encounters objects that are of different density or acoustic characteristics than the surrounding medium, such as fish, that reflect sound back toward the sound source. These echoes provide information on fish size, location, abundance and behavior. Data is usually processed and analysed using a variety of software such as Echoview.

Wave measurement

An upward looking echo sounder mounted on the bottom or on a platform may be used to make measurements of wave height and period. From this statistics of the surface conditions at a location can be derived.

Water velocity measurement

Special short range sonars have been developed to allow measurements of water velocity.

Bottom type assessment

Sonars have been developed that can be used to characterise the sea bottom into, for example, mud, sand, and gravel. Relatively simple sonars such as echo sounders can be promoted to seafloor classification systems via add-on modules, converting echo parameters into sediment type. Different algorithms exist, but they are all based on changes in the energy or shape of the reflected sounder pings. Advanced substrate classification analysis can be achieved using calibrated (scientific) echosounders and parametric or fuzzy-logic analysis of the acoustic data (See: Acoustic Seabed Classification)

Bottom topography measurement

Side-scan sonars can be used to derive maps of the topography of an area by moving the sonar across it just above the bottom. Low frequency sonars such as GLORIA have been used for continental shelf wide surveys while high frequency sonars are used for more detailed surveys of smaller areas.

Sub-bottom profiling

Powerful low frequency echo-sounders have been developed for providing profiles of the upper layers of the ocean bottom.

Synthetic aperture sonar

Various synthetic aperture sonars have been built in the laboratory and some have entered use in mine-hunting and search systems. An explanation of their operation is given in synthetic aperture sonar.

Parametric sonar

Parametric sources use the non-linearity of water to generate the difference frequency between two high frequencies. A virtual end-fire array is formed. Such a projector has advantages of broad bandwidth, narrow beamwidth, and when fully developed and carefully measured it has no obvious sidelobes. Its major disadvantage is very low efficiency of only a few percent. P.J. Westervelt's seminal 1963 JASA paper summarizes the trends involved.

Chapter-12

Acoustic Signature and Acoustic Doppler Current Profiler

Acoustic signature

Acoustic signature is used to describe a combination of acoustic emissions of ships and submarines.

Contributing factors

The acoustic signature is made up of a number of individual elements. These include:

- Machinery noise: noise generated by a ships engines, propeller shafts, fuel pumps, air conditioning systems, etc.
- Cavitation noise: noise generated by the creation of gas bubbles by the turning of a ship's propellers.
- Hydrodynamic noise: noise generated by the movement of water displaced by the hull of a moving vessel.

These emissions depend on a hull's dimensions, the installed machinery and ship's displacement. Therefore different ship classes will have different combinations of acoustic signals that together form a unique signature.

Targeting

Hydrophones and Sonar operating in passive mode can detect acoustic signals radiated by otherwise invisible submarines, and use these signals to target attacks.

Modern naval mines and torpedoes such as the CAPTOR mine can be programmed to distinguish the acoustic signatures of different vessels, leaving friendly vessels unmolested and attacking high-value targets when faced with multiple possible targets, e.g. distinguishing an aircraft carrier from its escorts.

Countermeasures

Warship designers aim to reduce the acoustic signature of ships and submarines just as much as they aim to reduce the radar cross sections and infra-red signals. For submarines, as a prime factor in how they can be detected the reduction of the acoustic signature is a primary goal.

The acoustic signature can be reduced by

- fitting of machinery with the best possible mechanical tolerances and designed to produce a minimum of noise.
- decoupling the machinery from the hull by mounting machinery on rubber mounting blocks.
- designing propellers to reduce cavitation, this led to the development of large slow turning propellers, today there is a preference now for pump-jet propulsors over propellers.
- the fitting of anechoic tiles to the hull, however ill fitting and loose anechoic tiles can themselves be a source of noise.
- hydrodynamic efficiency to minimise the perturbation of water.
- care in minimising protrusions from the hull.

Trimaran warships

The RV Triton

For a time the Royal Navy toyed with the idea of the trimaran hulled Future Surface Combatant. These would have had a very low acoustic signature. With three blade like hulls these ships would have cut through the water with a minimum of hydrodynamic noise. Radiated mechanical noise would also be minimised by using propulsors powered by a diesel-electric power plant; with the diesels being placed in the superstructure to mechanically isolate them from the water. This project got as far as the construction of the research ship RV Triton to test the principle of a large scale trimaran design.

Acoustic Doppler Current Profiler

Head of an ADCP with the four transducers

ADCP view ahead, mounted on an oceanographic device for long term measurements in the deep sea

An **Acoustic Doppler Current Profiler** (**ADCP or ADP**) is a sonar that attempts to produce a record of water current velocities for a range of depths. They are made of ceramic materials, and contain transducers, an amplifier, a receiver, a mixer, an oscillator, a clock, a temperature sensor, a compass, a pitch and roll sensor, and computer components to save the information collected. ADCPs can be configured in many ways: side-listening, into rivers and canals for long term continuous discharge measurements, downward-listening and mounted on boats for instantaneous surveys in the ocean or rivers, and mounted on moorings, or the seabed for long term current & wave studies. They can stay underwater for years at a time, and have a battery back for an energy source. The sonar is used for oceanography, estuary, river and stream flow measurement, and weather forecasting.

Components

Depending on the field application, an ADCP may use one or more ceramics, (or other piezo materials) for transducers, which work in water similar to directional loudspeakers in air. These transducers are aimed such that the sound pulse travels through the water in different, but known directions. As the sound energy leaves and arrives at the transducer face it is shifted in frequency, known as the Doppler effect, by the relative velocity of the water. As that sound energy is returned (echo) by scatterers in the water the sound may also be shifted in frequency if there is relative velocity of water to scatterer. Trigonometry, averaging and some critical assumptions are used to calculate the velocity of the group of echoing scatters in a volume of water. By repetitive sampling of the return echo, and by "gating" the return data in time, the ADCP can produce a "profile" of water currents over a range of depths. Phased array techniques are also used to aim the sound (acoustic) energy, allowing for economical production of smaller ADCPs to accommodate a range of frequencies from 38 kHz to several megahertz.

In addition to the transducers, an ADCP typically has an electronic amplifier, receiver, mixer, oscillator, accurate clock, temperature sensor, compass, pitch and roll sensor, analog-to-digital converters, memory, digital signal processor and instruction set. The analog-to-digital converters (ADCs) and digital signal processor (DSP) are used to sample the returning signal, determine the Doppler shift, and sample the compass and other sensors in order to calculate range and a velocity vector relative to a known orientation.

Performance

There are a number of factors that affect accuracy, resolution and profile range. Most notable are: absorption, spreading, speed of sound in water, bandwidth of the sound energy, signal strength of the transmitted pulse and echo, size of transducer, beam width (sic) of the energy pulse travelling through the water, frequency, and a host of limitations associated with the signal processing techniques and hardware, including clock/oscillator accuracy.

ADCPs can be self-contained and operate from batteries for many years under the sea or remotely in a river or stream. Some time later they are retrieved and the historical current data is transferred from the ADCP memory to a computer and displayed using a variety of graphical and text-based software to observe the water current profiles. Or ADCPs can be connected to RS232, RS422, RS485, SDI-12, USB, Ethernet, Fiber, Modbus (SCADA) connections to provide real time, "live", monitoring of their output.

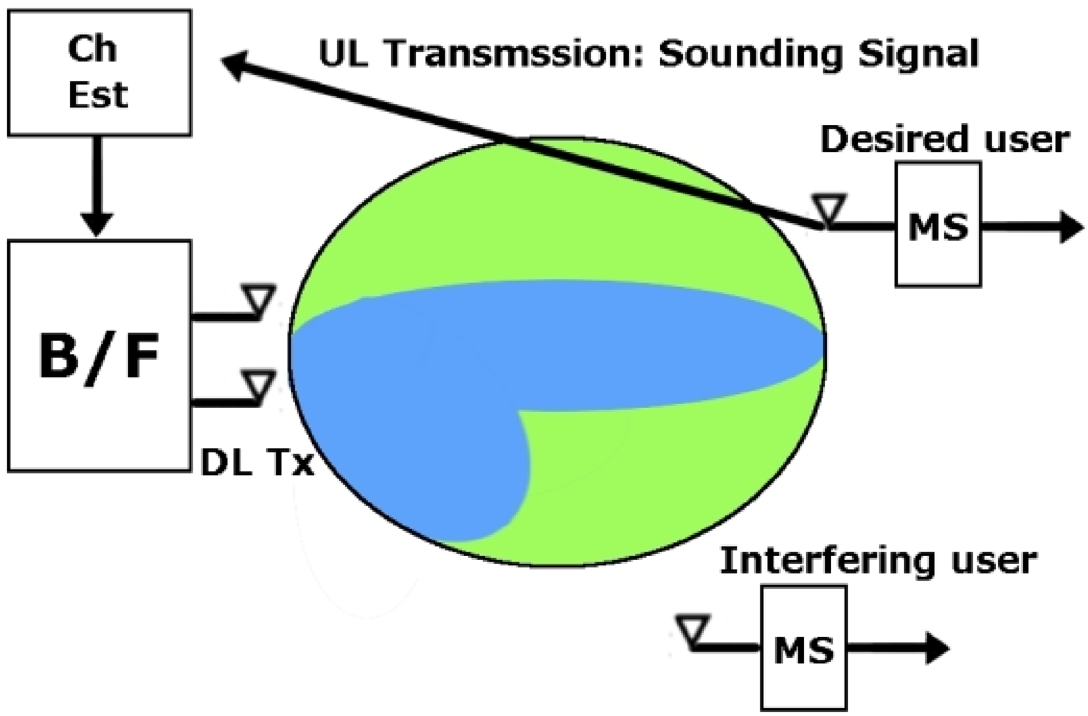
Uses

The ADCP, commercially available for about 25 years, is currently used for oceanography, estuary, river and stream flow measurement, even in weather forecasting. ADCPs are used in diverse ways, from locating underwater "tornadoes" that might damage deep water oil drilling activity, to measuring water flow through sewer pipes, or hanging up side down under an iceberg and measuring the flow of freshwater melting off the iceberg. Some harbor managers now use ADCPs to help them take advantage of tides and currents and optimize the flow of shipping in a busy port.

An ADCP can also be an acoustic **Doppler Velocity Log** (**DVL**) if it is programmed with the correct signal processing logic. The DVL bounces sound off the bottom (or a reference layer of water) and can determine the velocity vector of a subsea vehicle (or surface vessel) moving across the sea floor. This information can be combined with a starting fix, compass heading, and acceleration sensors (typically by use of a Kalman Filter) to calculate the position of the vehicle. DVLs are used to help navigate surface vessels, submarines, autonomous underwater vehicles, and ROVs for precise positioning in an environment where GPS, and other navigational aids, don't work.

Chapter-13

Beamforming



Beamforming

Beamforming is a signal processing technique used in sensor arrays for directional signal transmission or reception. This is achieved by combining elements in the array in such a way that signals at particular angle experience constructive interference and while others experience destructive interference. Beamforming can be used at both the transmit and receiver side to achieve spatially selectivity. The improvement compared with an omnidirectional reception/transmission is known as the receive/transmit gain (or loss).

Beamforming can be used for both radio or sound waves. It has found numerous applications in radar, sonar, seismology, wireless communications, radio astronomy, speech, acoustics, and biomedicine. Adaptive beamforming is used to detect and estimate the signal-of-interest at the output of a sensor array by means of data-adaptive spatial filtering and interference rejection.

Beamforming techniques

Beamforming takes advantage of interference to change the directionality of the array. When transmitting, a beamformer controls the phase and relative amplitude of the signal at each transmitter, in order to create a pattern of constructive and destructive interference in the wavefront. When receiving, information from different sensors is combined in such a way that the expected pattern of radiation is preferentially observed.

For example in sonar, to send a sharp pulse of underwater sound towards a ship in the distance, simply transmitting that sharp pulse from every sonar projector in an array simultaneously fails because the ship will first hear the pulse from the speaker that happens to be nearest the ship, then later pulses from speakers that happen to be the further from the ship. The beamforming technique involves sending the pulse from each projector at slightly different times (the projector closest to the ship last), so that every pulse hits the ship at exactly the same time, producing the effect of a single strong pulse from a single powerful projector. The same thing can be carried out in air using loudspeakers, or in radar/radio using antennas.

In passive sonar, and in reception in active sonar, the beamforming technique involves combining delayed signals from each hydrophone at slightly different times (the hydrophone closest to the target will be combined after the longest delay), so that every signal reaches the output at exactly the same time, making one loud signal, as if the signal came from a single, very sensitive hydrophone. Receive beamforming can also be used with microphones or radar antennas.

With narrow-band systems the time delay is equivalent to a "phase shift", so in this case the array of antennas, each one shifted a slightly different amount, is called a phased array. A narrow band system, typical of radars, is one where the bandwidth is only a small fraction of the centre frequency. With wide band systems this approximation no longer holds, which is typical in sonars.

In the receive beamformer the signal from each antenna may be amplified by a different "weight." Different weighting patterns (e.g., Dolph-Chebyshev) can be used to achieve the desired sensitivity patterns. A main lobe is produced together with nulls and sidelobes. As well as controlling the main lobe width (the beam) and the sidelobe levels, the position of a null can be controlled. This is useful to ignore noise or jammers in one particular direction, while listening for events in other directions. A similar result can be obtained on transmission.

Beamforming techniques can be broadly divided into two categories:

- conventional (fixed or switched beam) beamformers
- adaptive beamformers or adaptive arrays
    - Desired signal maximization mode
    - Interference signal minimization or cancellation mode

Conventional beamformers use a fixed set of weightings and time-delays (or phasings) to combine the signals from the sensors in the array, primarily using only information about the location of the sensors in space and the wave directions of interest. In contrast, adaptive beamforming techniques generally combine this information with properties of the signals actually received by the array, typically to improve rejection of unwanted signals from other directions. This process may be carried out in either the time or the frequency domain.

As the name indicates, an adaptive beamformer is able to automatically adapt its response to different situations. Some criterion has to be set up to allow the adaption to proceed such as minimising the total noise output. Because of the variation of noise with frequency, in wide band systems it may be desirable to carry out the process in the frequency domain.

Beamforming can be computationally intensive. Sonar phased array has a data rate low enough that it can be processed in real-time in software, which is flexible enough to transmit and/or receive in several directions at once. In contrast, radar phased array has a data rate so high that it usually requires dedicated hardware processing, which is hard-wired to transmit and/or receive in only one direction at a time. However, newer field programmable gate arrays are fast enough to handle radar data in real-time, and can be quickly re-programmed like software, blurring the hardware/software distinction.

Sonar beamforming requirements

Sonar itself has many applications, such as wide-area-search-and-ranging, underwater imaging sonars such as side-scan sonar and acoustic cameras.

Sonar beamforming implementation is similar in general technique but varies significantly in detail compared to electromagnetic system beamforming implementation. Sonar applications vary from 1 Hz to as high as 2 MHz, and array elements may be few and large, or number in the hundreds yet very small. This will shift sonar beamforming design efforts significantly between demands of such system components as the "front end" (transducers, preamps and digitizers) and the actual beamformer computational hardware downstream. High frequency, focused beam, multi-element imaging-search sonars and acoustic cameras often implement fifth-order spatial processing that places strains equivalent to Aegis radar demands on the processors.

Many sonar systems, such as on torpedoes, are made up of arrays of up to 100 elements that must accomplish beamsteering over a 100 degree field of view and work in both active and passive modes.

Sonar arrays are used both actively and passively in 1, 2, and 3 dimensional arrays.

- 1 dimensional "line" arrays are usually in multi-element passive systems towed behind ships and in single or multi-element side scan sonar.
- 2 dimensional "planar" arrays are common in active/passive ship hull mounted sonars and some side-scan sonar.
- 3 dimensional spherical and cylindrical arrays are used in 'sonar domes' in the modern submarine and ships.

Sonar differs from radar in that in some applications such as wide-area-search all directions often need to be listened to, and in some applications broadcast to, simultaneously. Thus a multibeam system is needed. In a narrowband sonar receiver the phases for each beam can be manipulated entirely by signal processing software, as compared to present radar systems that use hardware to 'listen' in a single direction at a time.

Sonar also uses beamforming to compensate for the significant problem of the slower propagation speed of sound as compared to that of electromagnetic radiation. In side-look-sonars, the speed of the towing system or vehicle carrying the sonar is moving at sufficient speed to move the sonar out of the field of the returning sound "ping". In addition to focusing algorithms intended to improve reception, many side scan sonars also employ beam steering to look forward and backward to "catch" incoming pulses that would have been missed by a single sidelooking beam.

Beamforming schemes

- A conventional beamformer can be a simple beamformer also known as delay-and-sum beamformer. All the weights of the antenna elements can have equal magnitudes. The beamformer is steered to a specified direction only by selecting appropriate phases for each antenna. If the noise is uncorrelated and there are no directional interferences, the signal-to-noise ratio of a beamformer with *L* antennas receiving a signal of power *P* is $\frac{1}{\sigma_n^2} P \cdot L$, where σ_n^2 is Noise variance or Noise power.
- Null-steering beamformer
- Frequency domain beamformer

Beamforming history in cellular standards

Beamforming techniques used in cellular phone standards have advanced through the generations to make use of more complex systems to achieve higher density cells, with higher throughput.

- Passive mode: (almost) non-standardized solutions
  - Wideband Code Division Multiple Access (WCDMA) supports direction of arrival (DOA) based beamforming
- Active mode: mandatory standardized solutions
  - 2G — Transmit antenna selection as an elementary beamforming
  - 3G — WCDMA: Transmit antenna array (TxAA) beamforming
  - 3G evolution — LTE/UMB: Multiple-input multiple-output (MIMO) precoding based beamforming with partial Space-Division Multiple Access (SDMA)
  - Beyond 3G (4G, 5G, …) — More advanced beamforming solutions to support SDMA such as closed loop beamforming and multi-dimensional beamforming are expected

Chapter-14

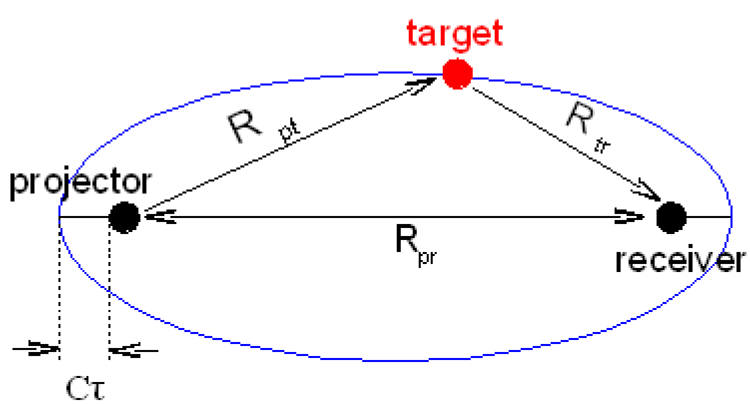
Bistatic Sonar

Most sonar systems are monostatic, in that the transmitter and receiver are in the same place. **Bistatic sonar** describes when the transmitter and receiver(s) are separated by a distance large enough to be comparable to the distance to the target.

Bistatic vs monostatic

Propagation (transmission) loss

This is a loss in sound level which happens while the sound pulse travels from projector to target and from target to receiver. There are 3 different mechanisms causing transmission loss: spherical (or cylindrical in shallow water) spreading, absorbing and scattering by ocean media inhomogeneities. Transmission loss (TL) is proportional to range, (the longer the sound travels the more the loss), and to sound frequency. In monostatic sonar the sound first travels from projector to target, then the same way back from target to receiver, so two-way loss is just 2TL, where TL is one-way loss. In bistatic sonar the total loss (in decibels) is a sum of TL_{pt} (from projector to target) and TL_{tr} (from target to receiver).

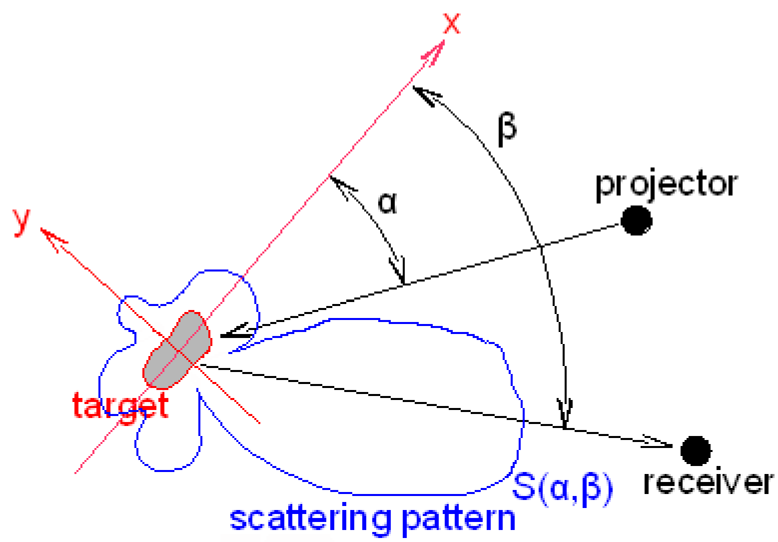


Bistatic sonar dead zone

Dead zone

In monostatic sonar, the first thing the receiver can hear is the sound of the transmitted ping. This sound level is very high, and it is impossible to detect the echo during the ping duration τ. That means targets are undetectable within the circle of $C\tau/2$ radius, where C is sound speed in water. This area is usually referred to as "dead zone". If the sonar is close to the surface, bottom or both, (which may happen in shallow water), the dead zone may be greater than $C\tau/2$ due to a high level of reverberation.

In bistatic sonar, the travel distance from projector to target and from target to receiver is $R = R_{pt} + R_{tr}$. As the projector is separated from receiver by R_{pr} distance, first R_{pr} /C seconds after the ping starts, the receiver is just waiting. After that time, it receives direct signal from the projector (often referred to as "direct blast",) which lasts $C\tau$ seconds. So the sonar cannot detect targets within the ellipse $R = Rpr + C\tau$, as shown at the picture. High level reverberation in the projector area does not affect the dead zone.



Target scattering pattern

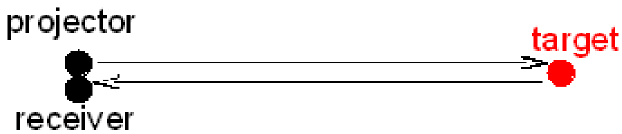
Target scattering pattern

Targets do not reflect the sound omni-directionally. The mechanism of sound reflection (or scattering by the target) is complicated, because the target is not just a rigid sphere. Scattered sound level depends on the angle β from which the target is ensonified by the projector, and it also varies with angle scattering direction α (refer to local target axes Z{x,y}). These angles are often referred to as aspects. This scattered sound level vs (α, β) function is called the scattering pattern S(α, β). Direction of maximum echo (maximum of S(α, β)) also depends on target shape and inner structure. So sometimes the best ensonifying aspect is not the same as the best receive aspect.

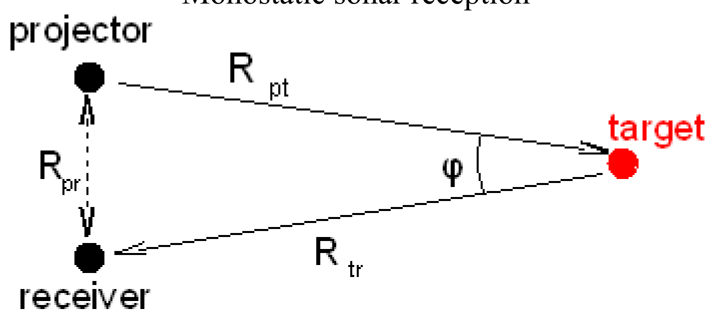
This leads to the bistatic solution. Target scattering becomes even more complicated if the target is buried (or semi-buried) into sea bottom sediments. (That happens to sea mines, waste containers, shipwrecks, etc.) In that case, the scattering mechanism is effected not by target features only, but also by sound wave interaction between the target and surrounding bottom.

Specific classes of bistatic sonars

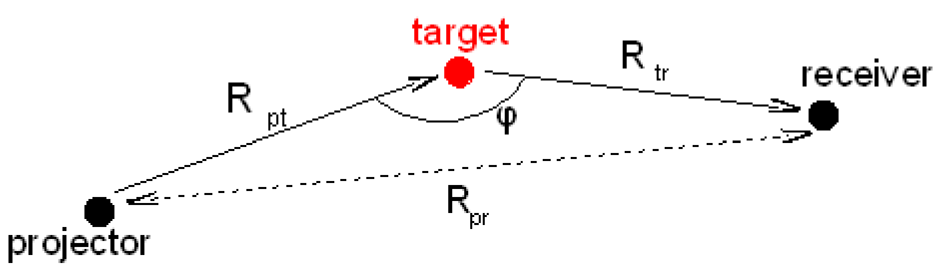
Backscattering and forward scattering



Monostatic sonar reception



Bistatic backscattering



Bistatic forward scattering

In **monostatic sonar** the receiver is listening to the echo which is reflected (scattered) right back from the target. Bistatic sonar can work in two ways: by utilizing either the target backscattering or forward scattering. **Backscattering** bistatic sonar is the sonar in which the bistatic angle φ is less than 90º. **Forward scattering** is the physical phenomena based on Babinet's principle. Forward scattering bistatic sonar is the sonar in which the bistatic angle φ is greater than 90º.

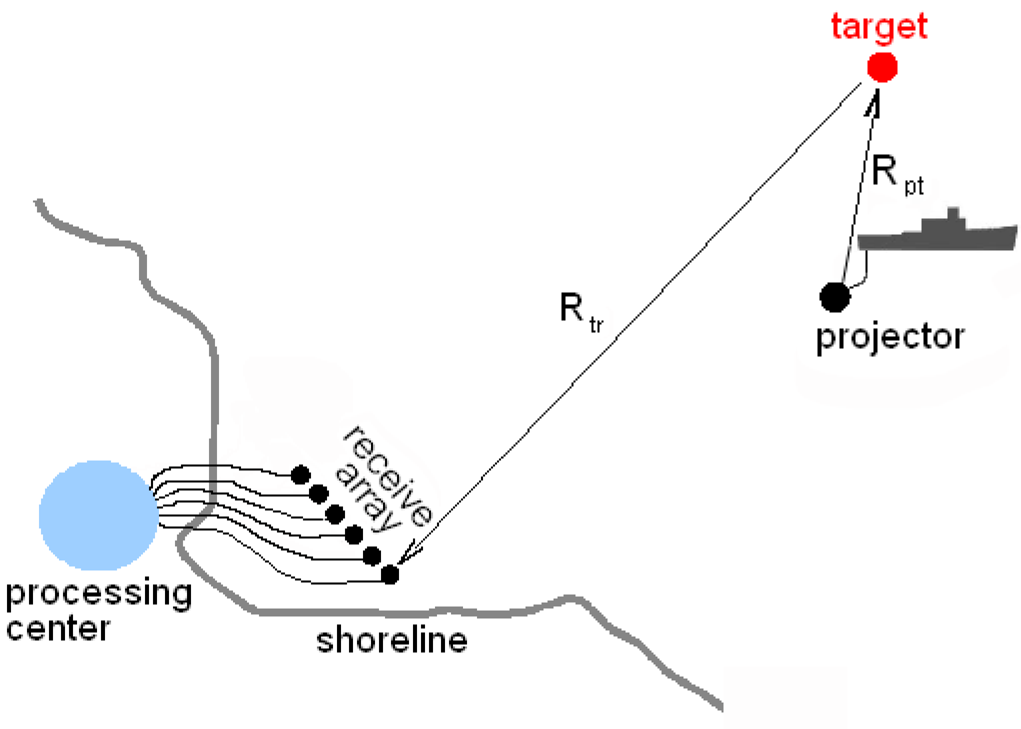
Pseudo-monostatic sonar

This is the sonar with a small bistatic angle. In other words, both the range from projector to target R_{pt} and from target to receiver R_{tr} is much greater than the distance from projector to receiver R_{pr}.

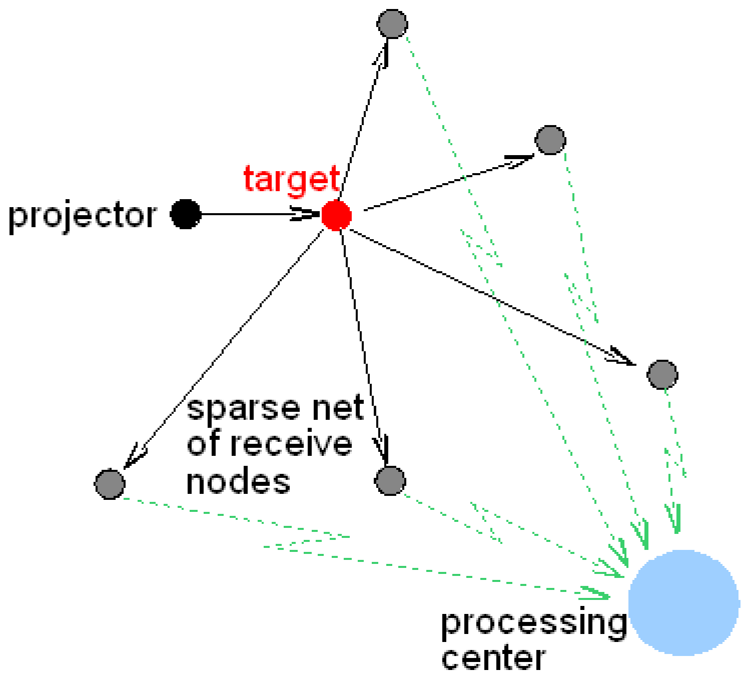
Multistatic sonar

This is the multi-node system with more than one projector, receiver or both.

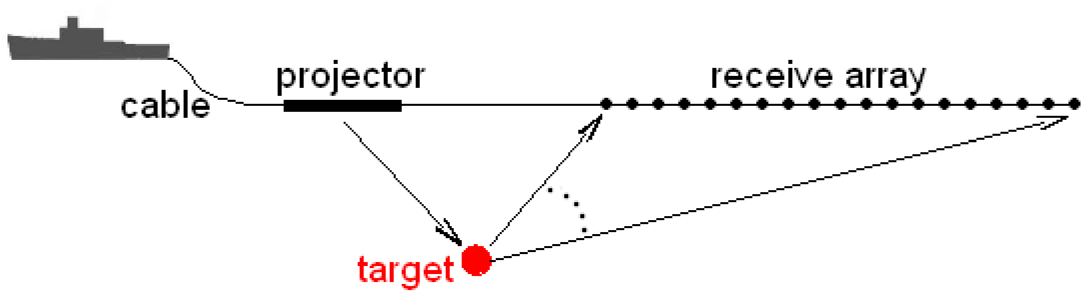
Applications



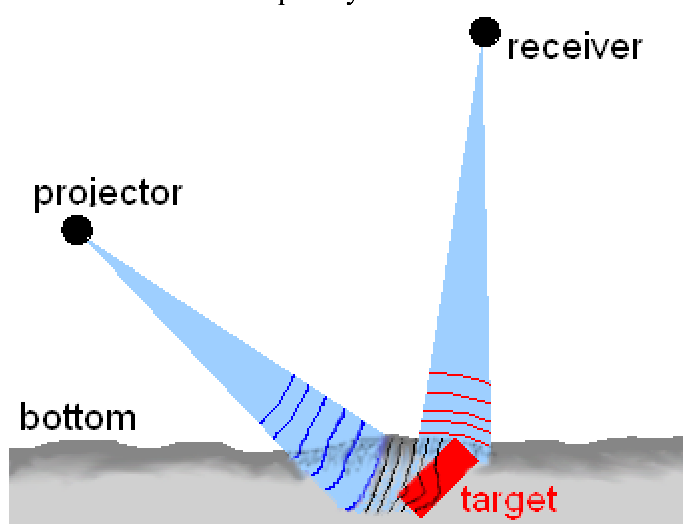
Long range surveillance



A net of receivers with a single projector



Low frequency towed sonar



Buried objects detection

Long range surveillance

For coastal surveillance, a large receive array of hydrophones is usually deployed close to the shore and connected with cables to a land-based processing center. To enable long range target detection (far away from the shore), one can use a powerful mobile projector, deployable from the ship. A system of this kind exploits the idea of “bringing the projector closer to area of interest and getting the transmission loss down”.

Large area surveillance with a single projector and a net of receivers

A system of this type is multistatic. It exploits the idea of “cover the area of interest with a sparse net of receivers and ensonify the whole area with a powerful projector”. Receive nodes may be sonobuoys (with radio communication link to a processing center) or autonomous underwater vehicles (AUVs) with an acoustic communication link. The example is GOATS project, using AUVs as receive nodes.

Low frequency towed sonar

The lower the frequency, the less the transmission loss absorbing and scattering components. On the other hand, the lower the frequency, the larger the size of directional projector and receive array. So the ship-deployable long range sonar is a low frequency bistatic towed array sonar with spatially separated projector and receive array. The example is LFATS towed sonar.

Buried object detection

To detect a buried object, the transmit ping must penetrate into the bottom. That requires a powerful and highly directional projector. Next, a directional receiver should be placed at the point where the “target + surrounding bottom” reflection is the best. This is a bistatic system. The example is SITAR project, developed to find objects like toxic waste containers and mines.

The principal advantages of bistatic and multistatic sonar include:

- Lower procurement and maintenance costs (if using a third party's transmitter)
- Operation without a frequency clearance (if using a third party's transmitter)
- Covert operation of the receiver
- Increased resilience to electronic countermeasures as waveform being used and receiver location are potentially unknown
- Possible enhanced radar cross section of the target due to geometrical effects

The principal disadvantages of bistatic and multistatic sonar include:

- System complexity
- Costs of providing communication between sites
- Lack of control over transmitter (if exploiting a third party transmitter)

- Harder to deploy
- Reduced low-level coverage due to the need for line-of-sight from several locations

Chapter-15

Fessenden Oscillator, Bistatic Range and Long Base Line Sonar

Fessenden oscillator

A **Fessenden oscillator** is an electro-acoustic transducer invented by Reginald Fessenden starting in 1912 in association with the Submarine Signaling Company of Boston. It was the first successful sonar device . It has been supplanted by piezoelectric devices.

It was an early kind of hydrophone, using electromagnetic forces; similar in operating principle to a voice coil loudspeaker, it was adapted to work in water instead of air. It was capable of creating underwater sounds and of picking up their echo.

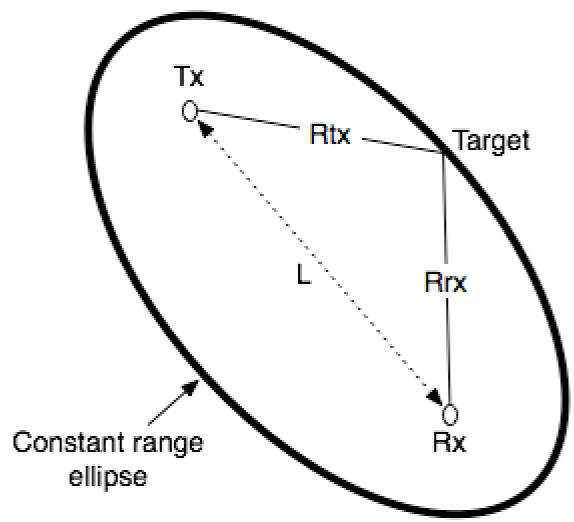
The creation of this device was motivated by a strongly-felt need to protect ships from collisions with obstacles and other ships.

Oscillator

The *oscillator* in the name referred to the fact that the device vibrated and moved water in response to a driving AC current. It was not an oscillator in the electronic sense that it generated a repetitive signal, in fact electronic oscillators did not yet exist when this device was created. Since the design of the device does not depend on a resonant response, it should not be considered a harmonic oscillator either.

Bistatic range

Bistatic range refers to the basic measurement of range made by a radar or sonar system with separated transmitter and receiver. The receiver measures the time difference of arrival of the signal from the transmitter directly, and via reflection from the target. This defines an ellipse of constant bistatic range, called an iso-range contour, on which the target lies, with foci centred on the transmitter and receiver. If the target is at range R_{rx} from the receiver and range R_{tx} from the transmitter, and the receiver and transmitter are a distance L apart, then the bistatic range is $R_{rx}+R_{tx}-L$. Motion of the target causes a rate of change of bistatic range, which results in bistatic Doppler shift.



Bistatic range geometry

Iso-range contour

Generally speaking, constant bistatic range points draw an ellipsoid with the transmitter and receiver positions as the focal points. The iso-range contours are where the ground slices the ellipsoid. When the ground is flat, this intercept forms an ellipse. Note that except when the two platforms have equal altitude, these ellipses are not centered on the specular point.

Long base line sonar

Long base line sonar, commonly referred to as **LBL**, is a method of acoustic positioning commonly used in deep water (water depth of greater than 3000 feet). A typical LBL positioning system consists of a transceiver and several beacons arranged into a structure called an array. The LBL transceiver pings each beacon and uses the 2-way travel time to calculate its position within the array. LBL positioning is much more accurate than ultra-short baseline (USBL) or SSBL surface-based positioning.

Advantages over surface tracking

In very deep water (1000m or more), acoustic positioning performed from the surface is subject to a large amount of noise and unusually long reply times due to the great distances the sound must travel. Also, a USBL system uses a single transceiver with multiple elements located close together and calculates range/bearings based on signal phase offsets. Because surface tracking may have an error radius of up to one-half percent of the water depth, 8,000 feet (2,400 m) of water can produce up to 40 feet (12 m) of error in any direction. The beacons in an LBL array are typically less than a kilometer apart, and the noise levels near the seabed in deep water are much less than near the surface, so LBL can usually resolve positions with less than a foot of error, regardless of the water depth. LBL combines range information from multiple sources. This redundancy helps to eliminate noise.

Beacon Array

Generally between 5 and 25 beacons will be placed on the seafloor. The acoustic navigator will then use a method of surface tracking such as USBL to lock in the locations of each beacon, which involves taking tens to hundreds of range/bearing measurements from different locations and averaging them to produce a final position. Once the position of each beacon is known to a high precision, the array is ready for navigation.

Underwater "smart" beacons are able to sample 2-way travel times between each other, so even a large array of smart beacons may be calibrated with only 2 or 3 known positions. The remaining beacons' positions can be resolved by ranging between beacons, reducing the time required to calibrate the array.

Tracking

Once the array has been calibrated, an underwater vehicle or diver may equip a transceiver which will take ranges (based on 2-way acoustic travel time) to each beacon in the array, and send that data to a computer located on the surface to be processed. The computer uses the ranges from these known points to calculate a final position.

Chapter-16

Geophysical MASINT

Geophysical MASINT is a branch of Measurement and Signature Intelligence (MASINT) that involves phenomena transmitted through the earth (ground, water, atmosphere) and manmade structures including emitted or reflected sounds, pressure waves, vibrations, and magnetic field or ionosphere disturbances.

According to the United States Department of Defense, MASINT is technically derived intelligence (excluding traditional imagery IMINT and signals intelligence SIGINT) that – when collected, processed, and analyzed by dedicated MASINT systems – results in intelligence that detects, tracks, identifies, or describes the signatures (distinctive characteristics) of fixed or dynamic target sources. MASINT was recognized as a formal intelligence discipline in 1986. Another way to describe MASINT is a "non-literal" discipline. It feeds on a target's unintended emissive by-products, the "trails" - the spectral, chemical or RF that an object leaves behind. These trails form distinct signatures, which can be exploited as reliable discriminators to characterize specific events or disclose hidden targets."

As with many branches of MASINT, specific techniques may overlap with the six major conceptual disciplines of MASINT defined by the Center for MASINT Studies and Research, which divides MASINT into Electro-optical, Nuclear, Geophysical, Radar, Materials, and Radiofrequency disciplines.

Military Requirements

Geophysical sensors have a long history in conventional military and commercial applications, from weather prediction for sailing, to fish finding for commercial fisheries, to nuclear test ban verification. New challenges, however, keep emerging.

First-world military forces, opposing other conventional militaries, there is an assumption that if a target can be located, it can be destroyed. As a result, concealment and deception have taken on new criticality. "Stealth" low-observability aircraft have gotten much attention, and new surface ship designs feature observability reduction. Operating in the confusing littoral environment produces a great deal of concealing interference.

Of course, submariners feel they invented low observability, and others are simply learning from them. They know that going deep or at least ultraquiet, and hiding among natural features, makes them very hard to detect.

Two families of military applications, among many, represent new challenges against which geophysical MASINT can be tried.

Deeply Buried Structures

One of the easiest ways for nations to protect weapons of mass destruction, command posts, and other critical structures is to bury them deeply, perhaps enlarging natural caves or disused mines. Deep burial is not only a means of protection against physical attack, as even without the use of nuclear weapons, there are deeply penetrating precision guided bombs that can attack them. Deep burial, with appropriate concealment during construction, is a way to avoid the opponent's knowing the buried facility's position well enough to direct precision guided weapons against it.

Finding deeply buried structures, therefore, is a critical military requirement. The usual first step in finding a deep structure is IMINT, especially using hyperspectral IMINT sensors to help eliminate concealment. "Hyperspectral images can help reveal information not obtainable through other forms of imagery intelligence such as the moisture content of soil. This data can also help distinguish camouflage netting from natural foliage." Still, a facility dug under a busy city would be extremely hard to find during construction. When the opponent knows that it is suspected that a deeply buried facility exists, there can be a variety of decoys and lures, such as buried heat sources to confuse infrared sensors, or simply digging holes and covering them, with nothing inside.

MASINT using as acoustic, seismic, and magnetic sensors would appear to have promise, but the reality of these sensors is that they must be fairly close to the target. Magnetic Anomaly Detection (MAD) is used, in antisubmarine water, for final localization before attack. The existence of the submarine is usually established through passive listening and refined with directional passive sensors and active sonar.

Once these sensors, as well as HUMINT and other sources, have failed, there is promise, for surveying large areas and deeply concealed facilities, using gravitimetric sensors. Gravity sensors are a new field, but military requirements are making it important while the technology to do it is becoming possible.

Naval Operations in Shallow Water

Especially in today's "green water" and "brown water" naval applications, navies are looking at MASINT solutions to meet new challenges of operating in littoral areas of operations. This symposium found it useful to look at five technology areas, which are interesting to contrast to the generally accepted categories of MASINT: acoustics and geology and geodesy/sediments/transport, nonacoustical detection (biology/optics/chemistry), physical oceanography, coastal meteorology, and electromagnetic detection.

Although it is unlikely there will ever be another World War II-style opposed landing on a fortified beach, another aspect of the littoral is being able to react to opportunities for amphibious warfare. Detecting shallow-water and beach mines remains a challenge, since mine warfare is a deadly "poor man's weapon."

While initial landings from an offshore force would be from helicopters or tiltrotor aircraft, with air cushion vehicles bringing ashore larger equipment, eventually, traditional landing craft, portable causeways, or other equipment will be needed to bring heavy equipment across a beach. Shallow depth and natural underwater obstacles can block a beach as well as can shallow-water mines. Synthetic Aperture Radar (SAR), airborne laser detection and ranging (LIDAR)) and use of bioluminescence to detect wake trails around underwater obstacles all may help solve this challenge.

Moving onto and across the beach has its own challenges. Remotely operated vehicles may be able to map landing routes, and they, as well as LIDAR and multispectral imaging, may be able to detect shallow water. Once on the beach, the soil has to support heavy equipment. Techniques here include estimating soil type from multispectral imaging, or from an airdropped penetrometer that actually measures the loadbearing capacity of the surface.

Weather and Sea Intelligence MASINT

The science and art of weather prediction used the ideas of measurement and signatures to predict phenomena, long before there were any electronic sensors. Masters of sailing ships might have no more sophisticated instrument than a wetted finger raised to the wind, and the flapping of sails.

Weather information, in the normal course of military operations, has a major effect on tactics. High winds and low pressures can change artillery trajectories. High and low temperatures cause both people and equipment to require special protection. Aspects of weather, however, also can be measured and compared with signatures, to confirm or reject the findings of other sensors.

The state of the art is to fuse meteorological, oceanographic, and acoustic data in a variety of display modes. Temperature, salinity and sound speed can be displayed horizontally, vertically, or in three-dimensional perspective.

Predicting Weather based on Measurements and Signatures

While early sailors had no sensors beyond their five senses, the modern meteorologist has a wide range of geophysical and electro-optical measuring devices, operating on platforms from the bottom of the sea to deep space. Prediction based on these measurements are based on signatures of past weather events, a deep understanding of theory, and computational models.

Weather predictions can give significant negative intelligence, when the signature of some combat system is such that it can operate only under certain weather conditions. Weather has long been an extremely critical part of modern military operations, as when the decision to land at Normandy on June 6, rather than June 5, 1944 depended on Dwight D. Eisenhower's trust in his staff weather advisor, Group Captain James Martin Stagg. It is rarely understood that something as fast as a ballistic missile reentry vehicle, or as "smart" as a precision guided munition, can still be affected by winds in the target area.

As part of Unattended Ground Sensors,. The Remote Miniature Weather Station (RMWS), from System Innovations, is an air-droppable version with a lightweight, expendable and modular system with two components: a meteorological (MET) sensor and a ceilometer (cloud ceiling height) with limited MET. The basic MET system is surface-based and measures wind speed and direction, horizontal visibility, surface atmospheric pressure, air temperature and relative humidity. The ceilometer sensor determines cloud height and discreet cloud layers. The system provides near-real-time data capable of 24-hour operation for 60 days. The RMWS can also go in with US Air Force Special Operations combat weathermen

The man-portable version, brought in by combat weathermen, has an additional function, as remote miniature ceilometer. Designed to measure multiple layer cloud ceiling heights and then send that data via satellite communications link to an operator display, the system uses a Neodinum YAG (NdYAG), 4 megawatt non-eye safe laser. According to one weatherman, "We have to watch that one," he said. "Leaving it out there basically we're worried about civilian populace going out there and playing with it—firing the laser and there goes somebody's eye. There are two different units [to RMWS]. One has the laser and one doesn't. The basic difference is the one with the laser is going to give you cloud height."

Hydrographic Sensors

Hydrographic MASINT is subtly different from weather, in that it considers factors such as water temperature and salinity, biologic activities, and other factors that have a major effect on sensors and weapons used in shallow water. ASW equipment, especially acoustic performance depends on the season the specific coastal site. Water column conditions, such as temperature, salinity, and turbidity are more variable in shallow than deep water. Water depth will influence bottom bounce conditions, as will the material of the bottom. Seasonal water column conditions (particularly summer versus winter) are inherently more variable in shallow water than in deep water.

While much attention is given to shallow waters of the littoral, other areas have unique hydrographic characteristics.

- regional areas with fresh water eddies
- open ocean salinity fronts
- near ice floes
- under ice

A submarine tactical development activity observed, "Fresh water eddies exist in many areas of the world. As we have experienced recently in the Gulf of Mexico using the Tactical Oceanographic Monitoring System (TOMS), there exist very distinct surface ducts that causes the Submarine Fleet Mission Program Library (SFMPL) sonar prediction to be unreliable. Accurate bathythermic information is paramount and a precursor for accurate sonar predictions."

Temperature and Salinity

Critical to the prediction of sound, needed by active and passive MASINT systems operating in water is knowing the temperature and salinity at specific depths. Antisubmarine aircraft, ships, and submarines can release independent sensors that measure the water temperature at various depths. The water temperature is critically important in acoustic detections, as changes in water temperature at thermoclines can act as a "barrier" or "layer" to acoustic propagation. To hunt a submarine, which is aware of water temperature, the hunter must drop acoustic sensors below the thermocline.

Water conductivity is used as a surrogate marker for salinity. The current and most recently developed software, however, does not give information on suspended material in the water or bottom characteristics, both considered critical in shallow-water operations.

The US Navy does this by dropping expendable probes, which transmit to a recorder, of 1978-1980 vintage, the AN/BQH-7 for submarines and the AN/BQH-71 for surface ships. While the redesign of the late seventies did introduce digital logic, the devices kept hard-to-maintain analog recorders, and maintainability became critical by 1995. A project was begun to extend with COTS components, to result in the AN/BQH-7/7A EC-3. In 1994-5, the maintainability of the in-service units became critical.

Variables in selecting the appropriate probe include:

- Maximum depth sounded
- Speed of launching vessel
- Resolution Vertical Distance Between Data Points (ft)
- Depth Accuracy

Biomass

Large schools of fish contain enough entrapped air to conceal the sea floor, or manmade underwater vehicles and structures. Fishfinders, developed for commercial and recreational fishing, are specialized sonars that can identify acoustic reflections between the surface and the bottom. Variations on commercial equipment are apt to be needed, especially in littoral areas rich in marine life.

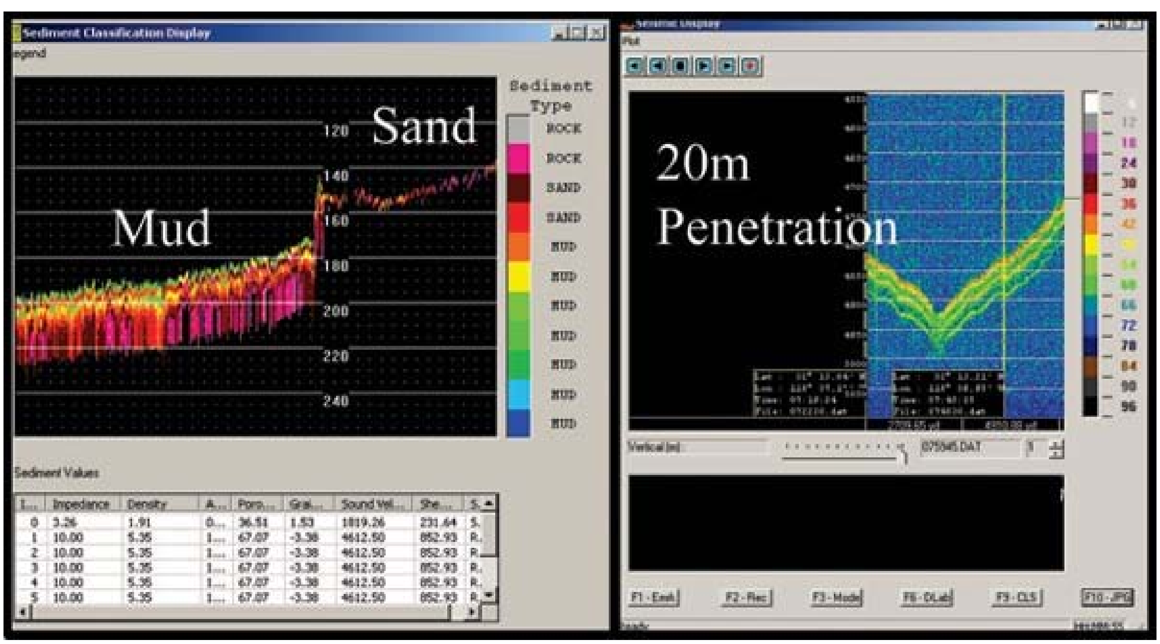
Sea Bottom Measurement

A variety of sensors can be used to characterise the sea bottom into, for example, mud, sand, and gravel. Active acoustic sensors are the most obvious, but there is potential information from gravitimetric sensors, electro-optical and radar sensors for making inferences from the water surface, etc.

Relatively simple sonars such as echo sounders can be promoted to seafloor classification systems via add-on modules, converting echo parameters into sediment type. Different algorithms exist, but they are all based on changes in the energy or shape of the reflected sounder pings.

Side-scan sonars can be used to derive maps of the topography of an area by moving the sonar across it just above the bottom. Multibeam hull-mounted sonars are not as precise as a sensor near the bottom, but both can give reasonable three-dimensional visualization.

Another approach comes from greater signal processing of existing military sensors. The US Naval Research Laboratory demonstrated both seafloor characterization, as well as subsurface characteristics of the seafloor. Sensors used, in different demonstrations, included normal incidence beams from the AM/UQN-4 surface ship depthfinder, and AN/BQN-17 submarine fathometer; backscatter from the Kongsberg EM-121 commercial multibeam sonar; AN/UQN-4 fathometers on mine countermeasures (MCM) ships, and the AN/AQS-20 mine-hunting system. These produced the "Bottom and Subsurface Characterization" graphic.



Bottom and Subsurface Characterization

Weather effects on chemical, biological, and radiological weapon propagation

One of the improvements in the Fuchs 2 reconnaissance vehicle is adding onboard weather instrumentations, including data such as wind direction and speed;, air and ground temperature; barometric pressure and humidity.

Acoustic MASINT

This includes the collection of passive or active emitted or reflected sounds, pressure waves or vibrations in the atmosphere (ACOUSTINT) or in the water (ACINT) or conducted through the ground Going well back into the Middle Ages, military engineers would listen to the ground for sounds of telltale digging under fortifications.

In modern times, acoustic sensors were first used in the air, as with artillery ranging in World War I. Passive hydrophones were used by the World War I Allies against German submarines; the UC-3, was sunk with the aid of hydrophone on 23 April 1916. Since submerged submarines cannot use radar, passive and active acoustic systems are their primary sensors. Especially for the passive sensors, the submarine acoustic sensor operators must have extensive libraries of acoustic signatures, to identify sources of sound.

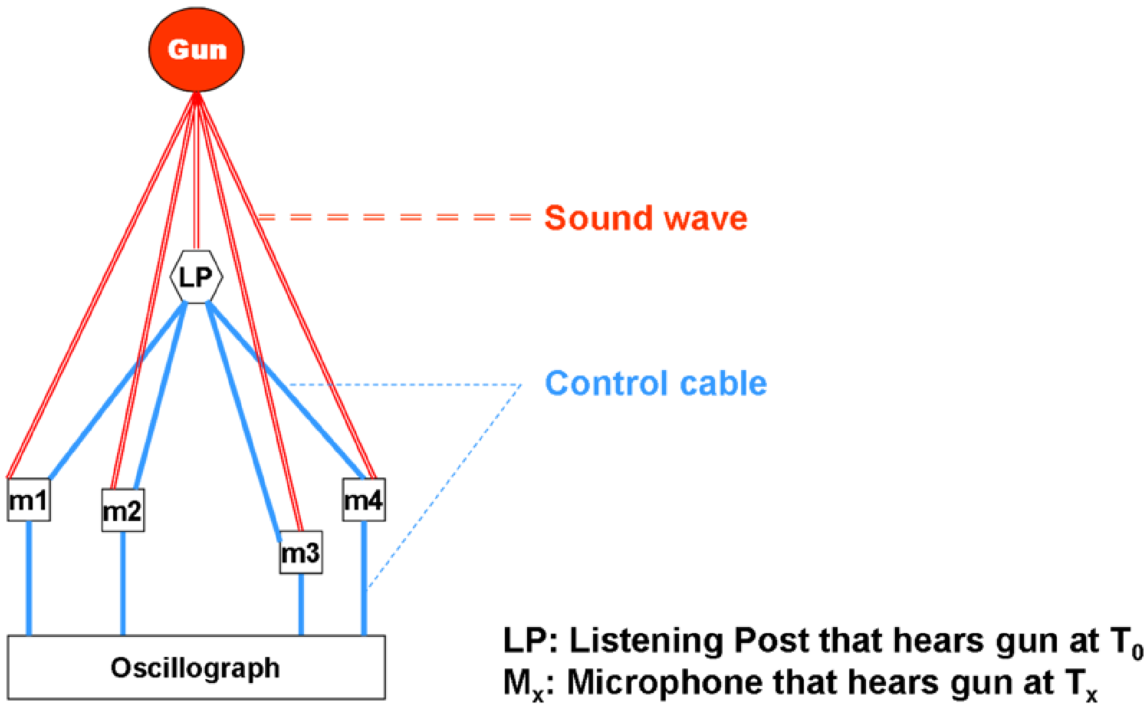
In shallow water, there are sufficient challenges to conventional acoustic sensors that additional MASINT sensors may be required. Two major confounding factors are:

- Boundary interactions. The effects of the seafloor and the sea surface on acoustic systems in shallow water are highly complex, making range predictions difficult. Multi-path degradation affects overall figure of merit and active classification. As a result, false target identifications are frequent.
- Practical limitations. Another key issue is the range dependence of shallow water propagation and reverberation. For example, shallow water limits the depth of towed sound detection arrays, thus increasing the possibility of the system's detecting its own noise. In addition, closer ship spacing increases the potential for mutual interference effects. It is believed that nonacoustic sensors, of magnetic, optical, bioluminescent, chemical, and hydrodynamic disturbances will be necessary in shallow-water naval operations.

Counterbattery and Countersniper Location and Ranging

While now primarily of historical interest, one of the first applications of acoustic and optical MASINT was locating enemy artillery by the sound and flash of their firing, a technique pioneered by Canadian Forces under Gen. Arthur Currie, with Andrew McNaughton in a key staff role. The combination of sound ranging (i.e., acoustic MASINT) and flash ranging (i.e., before modern optoelectronics) gave information unprecedented for the time, in both accuracy and timeliness. Enemy gun positions were located within 25 to 100 yards, with the information coming in three minutes or less.

Initial WWI Counterbattery Acoustic Systems



Sound Ranging

In the "Sound Ranging" graphic, the Listening Post, which is well forward of the microphone stations, sends an electrical signal to the microphone stations (MS) when the LP operator hears the gun's sound at time T_0. Either manually or electrically, each MS_x sends a starting pulse to an oscillograph. When the MS operator hears the sound, he stops the signal sent to the oscillograph. The oscillograph operator can then compute a time of arrival A_x, which is the difference between the T_0 and T_{Mx}. Without computer assistance, the range had to be computed manually.

The positions of the microphone stations and listening posts are precisely known. Each A_x can be graphed as a hyperbola. Where the asymptotes of the hyperbola meet is the position at which the gun is assumed to be located.

Where sound ranging is a time-of-arrival technique not dissimilar to that of modern multistatic sensors, flash ranging used theodolites to take bearings on the flash from the presurveyed flash observation post. The location of the gun was determined goniometrically, where the bearings intersected. Flash ranging, today, would be called electro-optical MASINT.

Artillery sound and flash ranging remained in use through World War II and into the postwar years, until mobile counterbattery radar, itself a MASINT radar sensor, became available. These techniques anticipated, and then paralleled, radio direction finding in SIGINT, which first was goniometric and now, with the precision time synchronization from GPS, is often time-of-arrival.

If the observation was at night, the Canadian master gunner could compare the sound and flash, while only sound was available in daylight. The Canadian units still had to estimate wind, temperature, and barometric pressure on the trajectory to the German artillery. and manually—and quickly—compute firing orders. Optimal times were on the order of three minutes.

Modern Acoustic Artillery Locators

Artillery positions now are located primarily with counterartillery radar, such as the US AN/TPQ-37, as well as IMINT. SIGINT also may give clues to positions, both with COMINT for firing orders, and ELINT for such things as weather radar. Still, there is renewed interest in both acoustic and electro-optical systems to complement counterartillery radar.

Acoustic sensors have come a long way since World War I. Typically, the acoustic sensor is part of a combined system, in which it cues radar or electro-optical sensors of greater precision, but narrower field of view.

HALO

The UK's hostile artillery locating system (HALO) has been in service with the British Army since 2002. HALO is not as precise as radar, but especially complements the directional radars. It passively detects artillery cannon, mortars and tank guns, with 360 degree coverage and can monitor over 2,000 square kilometers. HALO has worked in urban areas, the mountains of the Balkans, and the deserts of Iraq.

The system consists of a distributed array of up to 12 acoustic pressure sensors, which can compute location data on up to 8 rounds per second, and forwarding the data to the system operator. Assuming typical sensor dispersion, three or more sensors will measure the pressure wave, and the triangulation of the system computer can match a signature and help the AN/TPQ-36 and TPQ-37 Firefinder radars, which are not omnidirectional, to focus on the correct vector.

UTAMS

Another acoustic system is the US Army Unattended Transient Acoustic MASINT Sensor (UTAMS), which detects detect mortar and rocket launches and impacts. UTAMS has three to five acoustic arrays, each with four microphones, a processor, radio link, power source, and a laptop control computer. UTAMS, which was first operational in Iraq, first tested in November 2004 at a Special Forces Operating Base (SFOB) in Iraq. UTAMS was used in conjunction with AN/TPQ-36 and AN/TPQ-37 counter-artillery radar. While UTAMS was intended principally for detecting indirect artillery fire, Special

Forces and their fire support officer learned it could pinpoint improvised explosive device (IED) explosions and small arms/rocket-propelled grenade (RPG) fires. It detected Points of Origin (POO) up to 10 kilometers from the sensor.

Analyzing the UTAMS and radar logs revealed several patterns. The opposing force was firing 60 mm mortars during observed dining hours, presumably since that gave the largest groupings of personnel and the best chance of producing heavy casualties. That would have been obvious from the impact history alone, but these MASINT sensors established a pattern of the enemy firing locations.

This allowed the US forces to move mortars into range of the firing positions, give coordinates to cannon when the mortars were otherwise committed, and to use attack helicopters as a backup to both. The opponents changed to night fires, which, again, were countered with mortar, artillery, and helicopter fires. They then moved into an urban area where US artillery was not allowed to fire, but a combination of PSYOPS leaflet drops and deliberate near misses convinced the locals not to give sanctuary to the mortar crews.

Tower-mounted UTAMS array component of UTAMS in the Rocket Launch Spotter (RLS) system

Originally for a Marine requirement in Afghanistan, UTAMS was combined with electro-optical MASINT to produce the Rocket Launch Spotter (RLS) system useful against both rockets and mortars.

In the Rocket Launch Spotter (RLS) application, each array consists of four microphones and processing equipment. Analyzing the time delays between an acoustic wavefront's interaction with each microphone in the array UTAMS provides an azimuth of origin. The azimuth from each tower is reported to the UTAMS processor at the control station, and a POO is triangulated and displayed. The UTAMS subsystem can also detect and locate the point of impact (POI), but, due to the difference between the speeds of sound and light, it may take UTAMS as long as 30 seconds to determine the POO for a rocket launch 13 km away. In this application, the electro-optical component of RLS will detect the rocket POO earlier, while UTAMS may do better with the mortar prediction.

Passive sea-based acoustic sensors (hydrophones)

Modern hydrophones convert sound to electrical energy, which then can undergo additional signal processing, or that can be transmitted immediately to a receiving station. They may be directional or omnidirectional.

Navies use a variety of acoustic systems, especially passive, in antisubmarine warfare, both tactical and strategic. For tactical use, passive hydrophones, both on ships and airdropped sonobuoys, are used extensively in antisubmarine warfare. They can detect targets far further away than with active sonar, but generally will not have the precision location of active sonar, approximating it with a technique called Target Motion Analysis (TMA). Passive sonar has the advantage of not revealing the position of the sensor.

The Integrated Undersea Surveillance System (IUSS) consists of multiple subsystems in SOSUS, Fixed Distributed System (FDS), and the Advanced Deployable System (ADS or SURTASS). Reducing the emphasis on Cold War blue-water operations put SOSUS, with more flexible "tuna boat" sensing vessels called SURTASS being the primary blue-water long-range sensors

SURTASS used longer, more sensitive towed passive acoustic arrays than could be deployed from maneuvering vessels, such as submarines and destroyers.

USNS Able (T-AGOS-20) aft view of SURTASS equipment.

Air-dropped passive acoustic sensors

Passive sonobuoys, such as the AN/SSQ-53F, can be directional or omnidirectional and can be set to sink to a specific depth. These would be dropped from helicopters and maritime patrol aircraft such as the P-3.

Fixed underwater passive acoustic sensors

The US installed massive Fixed Surveillance System (FSS, also known as SOSUS) hydrophone arrays on the ocean floor, to track Soviet and other submarines.

Surface ship passive acoustic sensors

Purely from the standpoint of detection, towed hydrophone arrays offer a long baseline and exceptional measurement capability. Towed arrays, however, are not always feasible, because when deployed, their performance can suffer, or they can suffer outright damage, from fast speeds or radical turns. A state-of-the-art British towed array, with both passive and active capabilities, is Sonar 2087 made by Thales Underwater Systems.

Steerable sonar arrays on the hull or bow usually have a passive as well as active mode, as do variable-depth sonars

Surface ships may have warning receivers to detect hostile sonar.

Submarine passive acoustic sensors

Modern submarines have multiple passive hydrophone systems, such as a steerable array in a bow dome, fixed sensors along the sides of the submarines, and towed arrays. They also have specialized acoustic receivers, analogous to radar warning receivers, to alert the crew to the use of active sonar against their submarine.

US submarines made extensive clandestine patrols to measure the signatures of Soviet submarines and surface vessels. This acoustic MASINT mission included both routine patrols of attack submarines, and submarines sent to capture the signature of a specific vessel. US antisubmarine technicians on air, surface, and subsurface platforms had extensive libraries of vessel acoustic signatures.

Passive acoustic sensors can detect aircraft flying low over the sea.

Land-based Passive Acoustic Sensors (geophones)

Vietnam-era acoustic MASINT sensors included "Acoubuoy (36 inches long, 26 pounds) floated down by camouflaged parachute and caught in the trees, where it hung to listen. The Spikebuoy (66 inches long, 40 pounds) planted itself in the ground like a lawn dart. Only the antenna, which looked like the stalks of weeds, was left showing above ground." This was part of Operation Igloo White.

Part of the AN/GSQ-187 Improved Remote Battlefield Sensor System (I-REMBASS) is a passive acoustic sensor, which, with other MASINT sensors, detects vehicles and personnel on a battlefield. Passive acoustic sensors provide additional measurements that can be compared with signatures, and used to complement other sensors. I-REMBASS control will integrate, in approximately 2008, with the Prophet SIGINT/EW ground system.

For example, a ground search radar may not be able to differentiate between a tank and a truck moving at the same speed. Adding acoustic information, however, may quickly distinguish between them.

File:MASINT-T72-acoustic.jpg
The acoustic signature of a Russian T-72 tank shows the image is not of a dummy

Active Acoustic Sensors and Supporting Measurements

Combatant vessels, of course, made extensive use of active sonar, which is yet another acoustic MASINT sensor. Besides the obvious application in antisubmarine warfare, specialized active acoustic systems have roles in:

- Mapping the seafloor for navigation and collision avoidance. These include basic depth gauges, but quickly get into devices that do 3-dimensional underwater mapping
- Determining seafloor characteristics, for applications varying from understanding its sound-reflecting properties, to predicting the type of marine life that may be found there, to knowing when a surface is appropriate for anchoring or for using various equipment that will contact the seafloor

Various synthetic aperture sonars have been built in the laboratory and some have entered use in mine-hunting and search systems. An explanation of their operation is given in synthetic aperture sonar.

Water Surface, Fish Interference and Bottom Characterization

The water surface and bottom are reflecting and scattering boundaries. Large schools of fish, with air in their swim bladder balance apparatus, can also have a significant effect on acoustic propagation.

For many purposes, but not all naval tactical applications, the sea-air surface can be thought of as a perfect reflector. "The effects of the seafloor and the sea surface on acoustic systems in shallow water are highly complex, making range predictions difficult. Multi-path degradation affects overall figure of merit and active classification. As a result, false target identifications are frequent."

The acoustic impedance mismatch between water and the bottom is generally much less than at the surface and is more complex. It depends on the bottom material types and depth of the layers. Theories have been developed for predicting the sound propagation in the bottom in this case, for example by Biot and by Buckingham.

Water Surface

For high frequency sonars (above about 1 kHz) or when the sea is rough, some of the incident sound is scattered, and this is taken into account by assigning a reflection coefficient whose magnitude is less than one.

Rather than measuring surface effects directly from a ship, radar MASINT, in aircraft or satellites, may give better measurements. These measurements would then be transmitted to the vessel's acoustic signal processor.

Under Ice

A surface covered with ice, of course, is tremendously difference than even storm-driven water. Purely from a collision avoidance and acoustic propagation, a submarine needs to know how close it is to the bottom of ice. Less obvious is the need to know the three-dimensional structure of the ice, because submarines may need to break through it to launch missiles, to raise4 electronic masts, or to surface the boat. Three-dimensional ice information also can tell the submarine captain whether antisubmarine warfare aircraft can detect or attack the boat.

The state of the art is providing the submarine with a three-dimensional visualization of the ice above: the lowest part (ice keel) and the ice canopy. While sound will propagate differently in ice than liquid water, the ice still needs to be considered as a volume, to understand the nature of reverberations within it.

Bottom

A typical basic depth measuring device is the US AN/UQN-4A. Both the water surface and bottom are reflecting and scattering boundaries. For many purposes, but not all naval tactical applications, the sea-air surface can be thought of as a perfect reflector. In reality, there are complex interactions of water surface activity, seafloor characteristics, water temperature and salinity, and other factors that make "...range predictions difficult. Multi-path degradation affects overall figure of merit and active classification. As a result, false target identifications are frequent."

This device, however, does not give information on the characteristics of the bottom. In many respects, commercial fishing and marine scientists have equipment that is perceived as needed for shallow water operation.

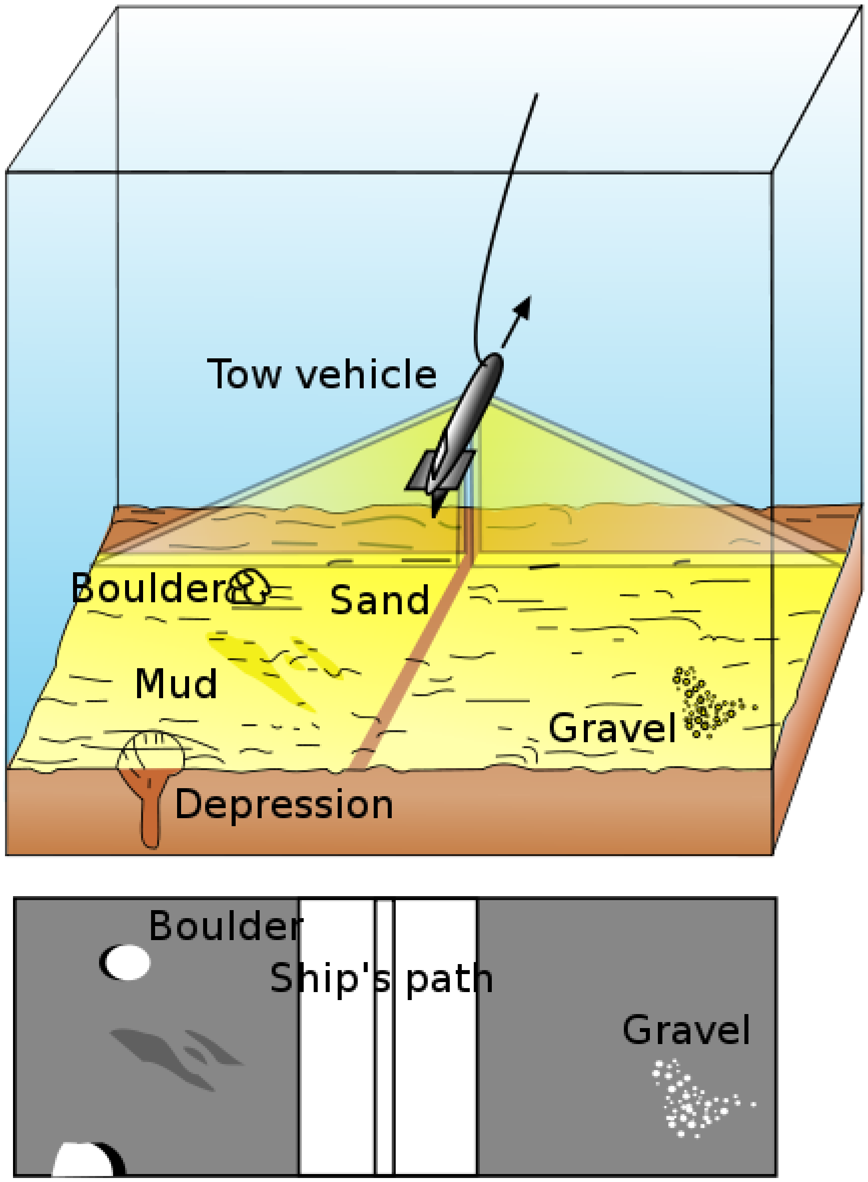


Diagram of sidescan sonar with towed probe, higher performance than multibeam ship-mounted but comparable

Biologic Effects on Sonar Reflection

A further complication is the presence of wind generated bubbles or fish close to the sea surface. . The bubbles can also form plumes that absorb some of the incident and scattered sound, and scatter some of the sound themselves. .

This problem is distinct from biologic interference caused by acoustic energy generated by marine life, such as the squeaks of porpoises and other cetaceans, and measured by acoustic receivers. The signatures of biologic sound generators need to be differentiated from more deadly denizens of the depths. Classifying biologics is a very good example of an acoustic MASINT process.

Surface Combatants

Modern surface combatants with an ASW mission will have a variety of active systems, with a hull- or bow-mounted array, protected from water by a rubber dome; a "variable-depth" dipping sonar on a cable, and, especially on smaller vessels, a fixed acoustic generator and receiver.

Some, but not all, vessels carry passive towed arrays, or combined active-passive arrays. These depend on target noise, which, in the combined littoral environment of ultraquiet submarines in the presence of much ambient noise. Vessels that have deployed towed arrays cannot make radical course maneuvers. Especially when active capabilities are included, the array can be treated as a bistatic or multistatic sensor, and act as a synthetic aperture sonar (SAS)

For ships that cooperate with aircraft, they will need a data link to sonobuoys and a sonobuoy signal processor, unless the aircraft has extensive processing capability and can send information that can be accepted directly by tactical computers and displays.

Signal processors not only analyze the signals, but constantly track propagation conditions. The former is usually considered part of a particular sonar, but the US Navy has a separate propagation predictor called the AN/UYQ-25B(V) Sonar *in situ* Mode Assessment System (SIMAS)

Echo Tracker Classifiers (ETC) are adjuncts, with a clear MASINT flavor, to existing surface ship sonars . ETC is an application of synthetic aperture sonar (SAS). SAS is already used for minehunting, but could help existing surface combatants, as well as future vessels and unmanned surface vehicles (USV), detect threats, such as very silent air-independent propulsion non-nuclear submarines, outside torpedo range. Torpedo range, especially in shallow water, is considered anything greater than 10 nmi.

Conventional active sonar may be more effective than towed arrays, but the small size of modern littoral submarines makes them difficult threats. Highly variable bottom paths, biologics, and other factors complicate sonar detection. If the target is slow-moving or waiting on the bottom, they have little or no Doppler effect, which current sonars use to recognize threats.

Continual active tracking measurement of all acoustically detected objects, with recognition of signatures as deviations from ambient noise, still gives a high false alarm rate (FAR) with conventional sonar. SAS processing, however, improves the resolution, especially of azimuth measurements, by assembling the data from multiple pings into a synthetic beam that gives the effect of a far larger receiver.

MASINT-oriented SAS measures shape characteristics and eliminates acoustically detected objects that do not conform to the signature of threats. Shape recognition is only one of the parts of the signature, which include course and Doppler when available.

Air-Dropped Active Sonobuoys

Active sonobuoys, containing a sonar transmitter and receiver, can be dropped from fixed-wing maritime patrol aircraft (e.g., P-3, Nimrod, Chinese Y-8, Russian and Indian Bear ASW variants), antisubmarine helicopters, and carrier-based antisubmarine aircraft (e.g., S-3). While there have been some efforts to use other aircraft simply as carriers of sonobuoys, the general assumption is that the sonobuoy-carrying aircraft can issue commands to the sonobuoys and receive, and to some extent process, their signals.

The Directional Hydrophone Command Activated Sonobuoy system (DICASS) both generate sound and listen for it. A typical modern active sonobuoy, such as the AN/SSQ 963D, generates multiple acoustic frequencies . Other active sonobuoys, such as the AN/SSQ 110B, generate small explosions as acoustic energy sources.

Airborne Dipping Sonar

Antisubmarine helicopters can carry a "dipping" sonar head at the end of a cable, which the helicopter can raise from or lower into the water. The helicopter would typically dip the sonar when trying to localize a target submarine, usually in cooperation with other ASW platforms or with sonobuoys. Typically, the helicopter would raise the head after dropping an ASW weapon, to avoid damaging the sensitive receiver. Not all variants of the same basic helicopter, even assigned to ASW, carry dipping sonar; some may trade the weight of the sonar for more sonobuoy or weapon capacity.

AN/AQS-13 Dipping sonar deployed from an H-3 Sea King used by numerous countries and produced in Italy, Japan, and the United Kingdom

The EH101 helicopter, used by a number of nations, has a variety of dipping sonars. The (British) Royal Navy version has Ferranti/Thomson-CSF sonar, while the Italian version uses the HELRAS. Russian Ka-25 helicopters carry dipping sonar, as does the US LAMPS SH-60 helicopter, which carries a "dipping" AQS-13F sonar, plus AN/SQQ-28(V)10 sonar signal processing for active sonobuoys it drops.

Surveillance Vessel Low-Frequency Active

Newer Low-Frequency Active (LFA) systems are controversial, as their very high sound pressures may be hazardous to whales and other marine life . A decision has been made to employ LFA on SURTASS vessels, after an environmental impact statement that indicated, if LFA is used with decreased power levels in certain high-risk areas for marine life, it would be safe when employed from a moving ship. The ship motion, and the variability of the LFA signal, would limit the exposure to individual sea animals. LFA operates in the low-frequency (LF) acoustic band of 100–500 Hz. It has an active component, the LFA proper, and the passive SURTASS hydrophone array. "The active component of the system, LFA, is a set of 18 LF acoustic transmitting source elements (called projectors) suspended by cable from underneath an oceanographic surveillance vessel, such as the Research Vessel (R/V) Cory Chouest, USNS Impeccable (T-AGOS 23), and the Victorious class (TAGOS 19 class).

"The source level of an individual projector is 215 dB. These projectors produce the active sonar signal or "ping." A "ping," or transmission, can last between 6 and 100 seconds. The time between transmissions is typically 6 to 15 minutes with an average transmission of 60 seconds. Average duty cycle (ratio of sound "on" time to total time) is less than 20 percent. The typical duty cycle, based on historical LFA operational parameters (2003 to 2007), is normally 7.5 to 10 percent."

This signal "...is not a continuous tone, but rather a transmission of waveforms that vary in frequency and duration. The duration of each continuous frequency sound transmission is normally 10 seconds or less. The signals are loud at the source, but levels diminish rapidly over the first kilometer."

Submarine Active Acoustic Sensors

The primary tactical active sonar of a submarine is usually in the bow, covered with a protective dome. Submarines for blue-water operations used active systems such as the AN/SQS-26 and AN/SQS-53 have been developed but were generally designed for convergence zone and single bottom bounce environments.

Submarines that operate in the Arctic also have specialized sonar for under-ice operation; think of an upside-down fathometer.

Submarines also may have minehunting sonar. Using measurements to differentiate between biologic signatures and signatures of objects that will permanently sink the submarine is as critical a MASINT application as could be imagined.

Active Acoustic Sensors for Minehunting

Sonars optimized to detect objects of the size and shapes of mines can be carried by submarines, remotely operated vehicles, surface vessels (often on a boom or cable) and specialized helicopters.

The classic emphasis on minesweeping, and detonating the mine released from its tether using gunfire, has been replaced with the AN/SLQ-48(V)2 mine neutralization system (MNS)AN/SLQ-48 - (remotely operated) Mine Neutralization Vehicle. This works well for rendering save mines in deep water, by placing explosive charges on the mine and/or its tether. The AN/SLQ-48 is not well suited to the neutralization of shallow-water mines. The vehicle tends to be underpowered and may leave on the bottom a mine that looks like a mine to any subsequent sonar search and an explosive charge subject to later detonation under proper impact conditions.

There is mine-hunting sonar, as well as (electro-optical) television on the ROV, and AN/SQQ-32 minehunting sonar on the ship.

Acoustic sensing of large explosions

An assortment of time-synchronized sensors can characterize conventional or nuclear explosions. One pilot study, the Active Radio Interferometer for Explosion Surveillance (ARIES). This technique implements an operational system for monitoring ionospheric pressure waves resulting from surface or atmospheric nuclear or chemical explosives. Explosions produce pressure waves that can be detected by measuring phase variations between signals generated by ground stations along two different paths to a satellite. This is a very modernized version, on a larger scale, of World War I sound ranging.

As can many sensors, ARIES can be used for additional purposes. Collaborations are being pursued with the Space Forecast Center to use ARIES data for total electron content measures on a global scale, and with the meteorology/global environment community to monitor global climate change (via tropospheric water vapor content measurements), and by the general ionospheric physics community to study travelling ionospheric disturbances.

Sensors relatively close to a nuclear event, or a high-explosive test simulating a nuclear event, can detect, using acoustic methods, the pressure produced by the blast. These include infrasound microbarographs (acoustic pressure sensors) that detect very low-frequency sound waves in the atmosphere produced by natural and man-made events.

Closely related to the microbarographs, but detecting pressure waves in water, are hydro-acoustic sensors, both underwater microphones and specialized seismic sensors that detect the motion of islands.

Seismic MASINT

US Army Field Manual 2-0 defines seismic intelligence as "The passive collection and measurement of seismic waves or vibrations in the earth surface." One strategic application of seismic intelligence makes use of the science of seismology to locate and characterize nuclear testing, especially underground testing. Seismic sensors also can characterize large conventional explosions that are used in testing the high-explosive components of nuclear weapons. Seismic intelligence also can help locate such things as large underground construction projects.

Since many areas of the world have a great deal of natural seismic activity, seismic MASINT is one of the emphatic arguments that there must be a long-term commitment to measuring, even during peacetime, so that the signatures of natural behavior is known before it is necessary to search for variations from signatures.

Strategic Seismic MASINT

For nuclear test detection, seismic intelligence is limited by the "threshold principle" coined in 1960 by George Kistiakowsky, which recognized that while detection technology would continue to improve, there would be a threshold below which small explosions could not be detected.

Tactical Seismic MASINT

The most common sensor in the Vietnam-era "McNamara Line" of remote sensors was the ADSID (Air-Delivered Seismic Intrusion Detector) sensed earth motion to detect people and vehicles. It resembled the Spikebuoy, except it was smaller and lighter (31 inches long, 25 pounds). The challenge for the seismic sensors (and for the analysts) was not so much in detecting the people and the trucks as it was in separating out the false alarms generated by wind, thunder, rain, earth tremors, and animals—especially frogs."

Vibration MASINT

This subdiscipline is also called **piezoelectric MASINT** after the sensor most often used to sense vibration, but vibration detectors need not be piezoelectric. Note that some discussions treat seismic and vibration sensors as a subset of acoustic MASINT. Other possible detectors could be moving coil or surface acoustic wave. . Vibration, as a form of geophysical energy to be sensed, has similarities to acoustic and seismic MASINT, but also has distinct differences that make it useful, especially in unattended ground sensors (UGS). In the UGS application, one advantage of a piezoelectric sensor is that it generates electricity when triggered, rather than consuming electricity, an important consideration for remote sensors whose lifetime may be determined by their battery capacity.

While acoustic signals at sea travel through water, on land, they can be assumed to come through the air. Vibration, however, is conducted through a solid medium on land. It has a higher frequency than is typical of seismic conducted signals.

A typical detector, the Thales MA2772 vibration is a piezoelectric cable, shallowly buried below the ground surface, and extended for 750 meters. Two variants are available, a high-sensitivity version for personnel detection, and lower-sensitivity version to detect vehicles. Using two or more sensors will determine the direction of travel, from the sequence in which the sensors trigger.

In addition to being buried, piezoelectric vibration detectors, in a cable form factor, also are used as part of high-security fencing. They can be embedded in walls or other structures that need protection.

Magnetic MASINT

A magnetometer is a scientific instrument used to measure the strength and/or direction of the magnetic field in the vicinity of the instrument. The measurements they make can be compared to signatures of vehicles on land, submarines underwater, and atmospheric radio propagation conditions. They come in two basic types:

- **scalar magnetometers** measure the total strength of the magnetic field to which they are subjected, and
- **vector magnetometers** have the capability to measure the component of the magnetic field in a particular direction.

Earth's magnetism varies from place to place and differences in the Earth's magnetic field (the magnetosphere) can be caused by two things:

- the differing nature of rocks
- the interaction between charged particles from the sun and the magnetosphere

Metal detectors use electromagnetic induction to detect metal. They can also determine the changes in existing magnetic fields caused by metallic objects.

Indicating Loops for detecting Submarines

One of the first means for detecting submerged submarines, first installed by the Royal Navy in 1914, was the effect of their passage over an anti-submarine indicator loop on the bottom of a body of water. A metal object passing over it, such as a submarine, will, even if degaussed, have enough magnetic properties to induce a current in the loop's cable. . In this case, the motion of the metal submarine across the indicating coil acts as an oscillator, producing electrical current.

MAD

A **magnetic anomaly detector** (MAD) is an instrument used to detect minute variations in the Earth's magnetic field. The term refers specifically to magnetometers used either by military forces to detect submarines (a mass of ferromagnetic material creates a detectable disturbance in the magnetic field)Magnetic anomaly detectors were first employed to detect submarines during World War II. MAD gear was used by both Japanese and U.S. anti-submarine forces, either towed by ship or mounted in aircraft to detect shallow submerged enemy submarines. After the war, the U.S. Navy continued to develop MAD gear as a parallel development with sonar detection technologies.

To reduce interference from electrical equipment or metal in the fuselage of the aircraft, the MAD sensor is placed at the end of a boom or a towed aerodynamic device. Even so, the submarine must be very near the aircraft's position and close to the sea surface for detection of the change or anomaly. The detection range is normally related to the distance between the sensor and the submarine. The size of the submarine and its hull composition determine the detection range. MAD devices are usually mounted on aircraft

MAD rear boom on P-3C

The SH-60B Seahawk helicopter carries an orange, towed MAD array known as a ‘MAD bird’, seen on the aft fuselage.

or helicopters.

There is some misunderstanding of the mechanism of detection of submarines in water using the MAD boom system. Magnetic moment displacement is ostensibly the main disturbance, yet submarines are detectable even when oriented parallel to the Earth's magnetic field, despite construction with non-ferromagnetic hulls.

For example, the Soviet-Russian Alfa class submarine, was constructed out of titanium. This light, strong material, as well as a unique nuclear power system, allowed the submarine to break speed and depth records for operational boats. It was thought that nonferrous titanium would defeat magnetic ASW sensors, but this was not the case. to give dramatic submerged performance and protection from detection by MAD sensors, is still detectable.

Since titanium structures are detectable, MAD sensors do not directly detect deviations in the Earth's magnetic field. Instead, they may be described as long-range electric and electromagnetic field detector arrays of great sensitivity.

An electric field is set up in conductors experiencing a variation in physical environmental conditions, providing that they are contiguous and possess sufficient mass. Particularly in submarine hulls, there is a measurable temperature difference between the bottom and top of the hull producing a related salinity difference, as salinity is affected by temperature of water. The difference in salinity creates an electric potential across the hull. An electric current then flows through the hull, between the laminae of sea-water separated by depth and temperature. The resulting dynamic electric field produces an electromagnetic field of its own, and thus even a titanium hull will be detectable on a MAD scope, as will a surface ship for the same reason.

Vehicle Detectors

The Remotely Emplaced Battlefield Surveillance System (REMBASS) is a US Army program for detecting the presence, speed, and direction of a ferrous object, such as a tank. Coupled with acoustic sensors that recognize the sound signature of a tank, it could offer high accuracy. It also collects weather information.

The Army's AN/GSQ-187 Improved Remote Battlefield Sensor System (I-REMBASS) includes both magnetic-only and combined passive infrared/magnetic intrusion detectors. The DT-561/GSQ hand emplaced MAG "sensor detects vehicles (tracked or wheeled) and personnel carrying ferrous metal. It also provides information on which to base a count of objects passing through its detection zone and reports their direction of travel relative to its location. The monitor uses two different (MAG and IR) sensors and their identification codes to determine direction of travel.

Magnetic detonators and countermeasures

Magnetic sensors, much more sophisticated than the early inductive loops, can trigger the explosion of mines or torpedoes. Early in World War II, the US tried to put magnetic torpedo exploder far beyond the limits of the technology of the time, and had to disable it, and then work on also-unreliable contact fuzing, to make torpedoes more than blunt objects than banged into hulls.

Since water is incompressible, an explosion under the keel of a vessel is far more destructive than one at the air-water interface. Torpedo and mine designers want to place the explosions in that vulnerable spot, and countermeasures designers want to hide the magnetic signature of a vessel. Signature is especially relevant here, as mines may be made selective for warships, merchant vessel unlikely to be hardened against underwater explosions, or submarines.

A basic countermeasure, started in World War II, was degaussing, but it is impossible to remove all magnetic properties.

Detecting landmines

Landmines often contain enough ferrous metal to be detectable with appropriate magnetic sensors. Sophisticated mines, however, may also sense a metal-detection oscillator, and, under preprogrammed conditions, detonate to deter demining personnel.

Foerster Minex 2FD 4.500 Metal detector used by the French army.

Not all landmines have enough metal to activate a magnetic detector. While, unfortunately, the greatest number of unmapped minefields are in parts of the world that cannot afford high technology, a variety of MASINT sensors could help demining. These would include ground-mapping radar, thermal and multispectral imaging, and perhaps synthetic aperture radar to detect disturbed soil.

Gravitimetric MASINT

While high school physics students are told that the value of gravity is 9.8 meters per second squared, they also learn Newton's equation that predicts that gravity is a function of mass. Given sufficiently sensitive instrumentation, it is possible to detect variations in gravity from the different densities of natural materials: the value of gravity will be greater on top of a granite monolith than over a sand beach. Again with sufficiently sensitive instrumentation, it should be possible to detect gravitational differences between solid rock, and rock excavated for a hidden facility.

Streland 2003 points out that the instrumentation indeed must be sensitive: variations of the force of gravity on the earth's surface are on the order of 10^6 of the average value. A practical gravitimetric detector of buried facilities would need to be able to measure "less than one one millionth of the force that caused the apple to fall on Sir Isaac Newton's head." To be practical, it would be necessary for the sensor to be able to be used while in motion, measuring the change in gravity between locations. This change over distance is called the *gravity gradient*, which can be measured with a gravity gradiometer.

Developing an operationally useful gravity gradiometer is a major technical challenge. One type, the SQUID Superconducting Quantum Interference Device gradiometer, may have adequate sensitivity, but it needs extreme cryogenic cooling, even if in space, a logistic nightmare. Another technique, far more operationally practical but lacking the necessary sensitivity, is the Gravity Recovery and Climate Experiment (GRACE) technique, currently using radar to measure the distance between pairs of satellites, whose orbits will change based on gravity. Substituting lasers for radar will make GRACE more sensitive, but probably not sensitive enough.

A more promising technique, although still in the laboratory, is quantum gradiometry, which is an extension of atomic clock techniques, much like those in GPS. Off-the-shelf atomic clocks measure changes in atomic waves over time rather than the spatial changes measured in a quantum gravity gradiometer. One advantage of using GRACE in satellites is that measurements can be made from a number of points over time, with a resulting improvement as seen in synthetic aperture radar and sonar. Still, finding deeply buried structures of human scale is a tougher problem than the initial goals of finding mineral deposits and ocean currents.

To make this operationally feasible, there would have to be a launcher to put fairly heavy satellites into polar orbits, and as many earth stations as possible to reduce the need for large on-board storage of the large amounts of data the sensors will produce. Finally, there needs to be a way to convert the measurements into a form that can be compared against available signatures in geodetic data bases. Those data bases would need significant improvement, from measured data, to become sufficiently precise that a buried facility signature would stand out.

Chapter-17

Marine Mammals and Sonar

A Humpback whale

Active sonar, the transmission equipment used on some ships to assist with navigation, has been suggested to be detrimental to the health and livelihood of some marine animals, although the precise mechanisms for this are not well understood. Some marine animals, such as whales and dolphins, use echolocation or "biosonar" systems to locate predators and prey. It is conjectured that active sonar transmitters could confuse these animals and interfere with basic biological functions such as feeding and mating.

History

The **SOFAR channel** (short for **sound fixing and ranging channel**), or **deep sound channel** (**DSC**), is a horizontal layer of water in the ocean centered around the depth at which the speed of sound is at a minimum. The SOFAR channel acts as a waveguide for sound, and low frequency sound waves within the channel may travel thousands of miles before dissipating. This phenomenon is an important factor in submarine warfare. The deep sound channel was discovered and described independently by Dr. Maurice Ewing, and Leonid Brekhovskikh in the 1940s.

Despite the use of the SOFAR channel in naval applications, the idea that animals might make use of this channel was not proposed until 1971. Roger Payne and Douglas Webb calculated that before ship traffic noise permeated the oceans, tones emitted by fin whales could have traveled as far as four thousand miles and still be heard against the normal background noise of the sea. Payne and Webb further determined that, on a quiet day in the pre–ship-propeller oceans, fin whale tones would only have fallen to the level of background noise after traveling thirteen thousand miles, that is, more than the diameter of the Earth.

Early confusion between fin whales and military sonar

Before extensive research on whale echolocation was completed, the low-frequency pulses emitted by some species of whales were often not correctly attributed to them. Dr Payne wrote:"Before it was shown that fin whales were the cause [of powerful sonar transmissions], no one could take seriously the idea that such regular, loud, low, and relatively pure frequency tones were coming from within the ocean, let alone from whales." This unknown sound was popularly known by navy acousticians as the *Jezebel Monster*. (*Jezebel* was narrow-band passive long-range sonar.) Some researchers believed that these sounds could be attributed to geophysical vibrations or an unknown Russian military program, and it wasn't until biologists William Schevill and William A. Watkins proved that whales possessed the biological capacity to emit sonar that the unknown sounds were correctly attributed.

Low frequency sonar

The electromagnetic spectrum has rigid definitions for "super low frequency", "extremely low frequency", "low frequency" and "medium frequency". Acoustics does not have a similar standard. The terms "low" and "mid" have roughly-defined historical meanings in sonar, because not many frequencies have been used over the decades. However, as more experimental sonars have been introduced, the terms have become muddled.

American low frequency sonar was originally introduced to the general public in a June 1961 Time magazine article, *New A.S.W. Artemis*, the low-frequency sonar used at the time, could fill a whole ocean with searching sound and spot anything sizable that was moving in the water. *Artemis* grew out of a 1951 suggestion by Harvard physicist Frederick V. Hunt (Artemis is the Ancient Greek goddess of the hunt), who convinced Navy anti-submarine experts that submarines could be detected at great distances only by unheard-of volumes of low-pitched sound. At the time, an entire *Artemis* system was envisioned to form a sort of underwater DEW (*Distant Early Warning*) line to warn the U.S. of hostile submarines. Giant, unattended transducers, powered by cables from land, would be lowered to considerable depths where sound travels best. The Time magazine article was published during the maiden voyage of the Soviet submarine *K-19*, which was the first Soviet submarine equipped with ballistic missiles. Four days later the submarine would have the accident that gave it its nickname. The impact on marine mammals by this system was certainly not a consideration. *Artemis* never became an operational system.

Low-frequency sonar was revived in the early 1980s for military and research applications. The idea that the sound could interfere with whale biologics became widely discussed outside of research circles when Scripps Institute of Oceanography borrowed and modified a military sonar for the Heard Island Feasibility Test conducted in January and February 1991. The sonar modified for the test was an early version of SURTASS deployed in the MV *Cory Chouest*. As a result of this test a "Committee on Low-Frequency Sound and Marine Mammals" was organized by the National Research Council. Their findings were published in 1994, in *Low-Frequency Sound and Marine Mammals: Current Knowledge and Research Needs*.

Long-range transmission does not require high power. All frequencies of sound lose an average of 65dB in the first few seconds before the sound waves strike the ocean bottom. After that the acoustic energy in mid or high-frequency sound is converted into heat, primarily by the epsom salt dissolved in sea water. Very little of low frequency acoustic energy is not converted into heat, so the signal can be detected for long ranges. Fewer than five of the transducers from the low frequency active array were used in the Heard Island Feasibility Test, and the sound was detected on the opposite side of the Earth. The transducers were temporarily altered for this test to transmit sound at 50 hertz, which is lower than their normal operating frequency.

A year after the Heard Island Feasibility Test a new low-frequency active sonar was installed in the *Cory Chouest* with 18 transducers instead of 10. An environmental impact statement was prepared for that system.

Mid frequency sonar

The term *mid frequency* sonar is usually used to refer to sonars that project sound in the 3 to 4 kilohertz (kHz) range. Ever since the launch of the USS *Nautilus* (SSN-571) on 17 January 1955 the US Navy knew it was only a matter of time until the other naval powers had their own nuclear submarines. The mid-frequency sonar was developed for anti-submarine warfare against these future boats. The standard post-WWII active sonars (which were usually above 7 kHz) had an insufficient range against this new threat. Active sonar went from a piece of equipment attached to a ship, to a piece of equipment that was central to the design of a ship. They are described in the same 1961 Time magazine article by the quote "*the latest shipboard sonar weighs 30 tons and consumes 1,600 times as much power as the standard postwar sonar*". A modern system produced by Lockheed Martin since the early 1980s is the AN/SQQ-89. On June 13, 2001, Lockheed Martin announced that it had delivered its 100th AN/SQQ-89 undersea warfare system to the U.S. Navy.

There was anecdotal evidence that mid-frequency sonar could have adverse effects on whales dating back to the days of whaling. The following story is recounted in a book published in 1995:

Mid Frequency Sonar and Whaling
Source: ***Among Whales*** **by Roger Payne (pg 258) Published 2 June 1995**
Another innovation by the whalers was the use of sonar to track whales they were pursuing underwater. But there was a problem; as the boat gained on the whale, the whale started exhaling while still submerged. This produced a cloud of bubbles in the water that reflected sound better than the whale did and made a false target (akin to what a pilot does when releasing metal chaff to create a false radar echo). I suspect that this behavior by whales was simply fortuitous since exhaling while still submerged is simply a means by which a whale can reduce the time it has to remain at the surface, where surface drag will slow it down. *Whalers quickly discovered that a frequency of* ***three thousand hertz*** *seemed to panic the whales, causing them to surface much more often for air, This was a "better" use for sonar because it afforded the whalers more chances to shoot the whales. So they equipped their catcher boats with sonar at that frequency. Of course the sonar also allows the whalers to follow the whale underwater, but that is its secondary use. Its primary use is for scaring whales so that they start "panting" at the surface.*

In 1996 twelve Cuvier's beaked whales beached themselves alive along the coast of Greece while NATO (North Atlantic Treaty Organisation) was testing an active sonar with combined low and mid-range frequency transducers, according to a paper published in the journal Nature in 1998. The author established for the first time the link between atypical mass strandings of whales and the use of military sonar by concluding that *although pure coincidence cannot be excluded* there was better than a 99.3% likelihood that sonar testing caused that stranding. He noted that the whales were spread along 38.2 kilometres of coast and were separated by a mean distance of 3.5 km (sd=2.8, n=11). This spread in time and location was atypical, as usually whales mass strand at the same place and at the same time.

At the time that Dr. Frantzis wrote the article he was unaware of several important factors.

- The time correlation was much tighter than he knew. He knew about the test from a notice to mariners which only published that the test would occur over a five day period within a large area of the ocean. In fact the first time the sonar was turned on was the morning of 12 May 1996, and six whales stranded that afternoon. The next day the sonar was turned on again and another six whales stranded that afternoon. Without knowing the coordinates of the ships he would not have realized that the ship was only about 10–15 miles offshore.

- The sonar being used in the test was an experimental research and development sonar, which was considerably smaller and less powerful than an operational sonar onboard a deployed naval vessel. Dr Frantzis believed that wide distribution of the stranded whales indicated that the cause has a large synchronous spatial extent and a sudden onset. Knowing that the sound source level was fairly low (it was only 226 dB (decibels) @ 3 kHz which is low compared to an operational sonar) would have made the damage mechanism even more puzzling.
- The experimental sonar used in the test, Towed Vertically Directive Source (TVDS) which had the dual 600 Hz and 3 kHz transducers, had been used for the first time in the Mediterranean Sea south of Sicily the year before in June 1995. Previous activated towed array sonar research using different sources on board the same ship included participation in NATO exercises "Dragon Hammer '92" and "Resolute Response '94".

Since the source level of this experimental sonar was only 226 dB @ 3 kHz re. 1 meter, at only 100 meters the received level would drop by 40 dB (to 186 dB). A NATO panel investigated the above stranding and concluded the whales were exposed to 150-160 dB re 1 μPa of low and mid-range frequency sonar. This level is about 66 dB less (more than a million times lower intensity) than the threshold for hearing damage specified by a panel of marine mammal experts.

The idea that a relatively low power sonar could cause a mass stranding of such a large number of whales was very unexpected by the scientific community. Most research had been focused on the possibility of masking signals, interference with mating calls, and similar biological functions. Deep diving marine mammals were species of concern, but very little definitive information was known. In 1995 a comprehensive book on the relation between marine mammals and noise had been published, and it did not even mention strandings.

Acoustically induced bubble formation

There was anecdotal evidence from whalers that sonar could panic whales and cause them to surface more frequently making them vulnerable to harpooning. It has also been theorized that military sonar may induce whales to panic and surface too rapidly leading to a form of decompression sickness. In general trauma caused by rapid changes of pressure is known as *barotrauma*. The idea of acoustically enhanced bubble formation was first raised by a paper published in *The Journal of the Acoustical Society of America* in 1996 and again *Nature* in 2003. It reported acute gas-bubble lesions (indicative of decompression sickness) in whales that beached shortly after the start of a military exercise off the Canary Islands in September 2002.

In the Bahamas in 2000, a sonar trial by the United States Navy of transmitters in the frequency range 3–8 kHz at a source level of 223–235 decibels re 1 μPa (scaled to a distance of 1 m) was associated with the beaching of seventeen whales, seven of which were found dead. Environmental groups claimed that some of the beached whales were bleeding from the eyes and ears, which they considered an indication of acoustically-induced trauma. The groups allege that the resulting disorientation may have led to the stranding.

Naval sonar-linked incidents

> “ Worldwide, use of active sonar has been linked to about 50 marine mammal strandings between 1996 and 2006. In all of these occurrences, there were other contributing factors, such as unusual (steep and complex) underwater geoqraphy, limited egress routes, and a specific species of marine mammal — beaked whales — that are suspected to be more sensitive to sound than other marine mammals. ”
>
> —Rear Admiral Lawrence Rice

Date	Location	Species and Number	Naval Activity
1963-05	Gulf of Genoa, Italy	Cuvier’s beaked whale (15) stranded	Naval maneuvers
1988-11	Canary Islands	Cuvier’s beaked whale (12+) Gervais' beaked whale (1) stranded	FLOTA 88 exercise
1989-10	Canary Islands	Cuvier's beaked whale (15+), Gervais' beaked whale (3), Blainville's beaked whale (2) stranded	CANAREX 89 exercise
1991-12	Canary Islands	Cuvier's beaked whale (2) stranded	SINKEX 91 exercise
1996-05-12	Gulf of Kyparissia, Greece	Cuvier's beaked whale (12) stranded	NATO Shallow Water Acoustic Classification exercise
1998-07	Kauai, Hawaii	beaked whale (1), sperm whale (1) stranded	RIMPAC 98 exercise
1999-10	U.S. Virgin Islands and Puerto Rico	Cuvier’s beaked whale (4) stranded	COMPTUEX exercise
2000-03-15	Bahamas	Cuvier’s beaked whale (9), Blainville’s beaked whale (3), beaked whale spp (2), Minke whale (2), Atlantic spotted dolphin (1) stranded	Naval MFA
2000-05-10	Madeira Island, Portugal	Cuvier’s beaked whale (3) stranded	NATO Linked Seas 2000 and MFA
2002-09	Canary Islands	Cuvier’s beaked whale (9), Gervais’ beaked whale (1), Blainville’s beaked whale (1), beaked whale spp. (3) stranded	Neo Tapon 2002 exercise and MFA
2003-05	Haro Strait, Washington	Harbor porpoise (14), Dall’s porpoise (1) Orca avoidance “stampede”	U.S.S. Shoup transiting while using MFA (AN/SQS-53C)
2004-07	Kauai, Hawaii	Melon-headed whale (~200) avoidance “stampede”	RIMPAC 04 exercise with MFA

2004-07-22	Canary Islands	Cuvier’s beaked whale (4) stranded	Majestic Eagle 04 exercise
2005-10-25	Marion Bay, Tasmania	Long-finned pilot whales (145) stranded	Two minesweepers using active sonar
2006-01-26	Almeira Coast, Spain	Cuvier's beaked whale (4) stranded	Royal Navy’s HMS Kent using active sonar
2008-06	Cornish coast, United Kingdom	Dolphins (26) stranded	Royal Navy sonar exercise

Court cases

Since mid-frequency sonar has been correlated with mass cetacean strandings throughout the world's oceans, it has been singled out by some environmentalists as a focus for activism. A lawsuit filed by the Natural Resources Defense Council in Santa Monica, California on 20 October 2005 contended that the U.S. Navy has conducted sonar exercises in violation of several environmental laws, including the National Environmental Policy Act, the Marine Mammal Protection Act, and the Endangered Species Act. Mid-frequency sonar is by far the most common type of active sonar in use by the world's navies, and has been widely deployed since the 1960s.

On November 13, 2007, a United States appeals court restored a ban on the U.S. Navy's use of submarine-hunting sonar in training missions off Southern California until it adopted better safeguards for whales, dolphins and other marine mammals. On 16 January 2008, President George W. Bush exempted the US Navy from the law and argued that naval exercises are crucial to national security. On 4 February 2008, a Federal judge ruled that despite President Bush's decision to exempt it, the Navy must follow environmental laws placing strict limits on mid-frequency sonar. In a 36-page decision, U.S. District Judge Florence-Marie Cooper wrote that the Navy is not "exempted from compliance with the National Environmental Policy Act" and the court injunction creating a 12-nautical-mile (22 km) no-sonar zone off Southern California. On 29 February 2008, a three-judge federal appeals court panel upheld the lower court order requiring the Navy to take precautions during sonar training to minimize harm to marine life. The U.S. Supreme Court overturned the ruling in a 5:4 decision on 12 November 2008. However, this does not spell the end of environmentalist challenges to the Navy's sonar activities. Indeed, environmentalists have already interpreted the *NRDC v. Winter* decision as narrowly as possible given the Court’s refusal to address their non-‘science’ claims.

Mitigation methods

Environmental impacts of the operation of active sonar are required to be carried out by US law. Procedures for minimising the impact of sonar are developed in each case where there is significant impact.

The impact of underwater sound can be reduced by limiting the sound exposure received by an animal. The maximum sound exposure level recommended by Southall et al. for cetaceans is 215 dB re 1 μPa^2 s for hearing damage. Maximum sound pressure level for behavioural effects is dependent on context (Southall et al.).

A great deal of the legal and media conflict on this issue has to do with questions of who determines what type of mitigation is sufficient. Coastal commissions, for example, were originally thought to only have legal responsibility for beachfront property, and state waters (three miles into sea). Because active sonar is instrumental to ship defense, mitigation measures that may seem sensible to a civilian agency without any military or scientific background can have disastrous effects on training and readiness. Navies therefore often define their own mitigation requirements.

Examples of mitigation measures include:

1. not operating at nighttime
2. not operating at specific areas of the ocean that are considered sensitive
3. slow ramp-up of intensity of signal to give whales a warning
4. air cover to search for mammals
5. not operating when a mammal is known to be within a certain range
6. onboard observers from civilian groups
7. using fish-finders to look for whales in the vicinity
8. large margins of safety for exposure levels
9. not operating when dolphins are bow-riding
10. operations at less than full power
11. paid teams of veterans to investigate strandings after sonar operation.

Besides the expense of some of the mitigation measures some of them might interfere with operations. For this reason, mitigation requirements for wartime use of naval sonars can differ from either civilian mitigation requirements or from military requirements during exercises or peacetime operations. Prohibition on night time operations may be a huge waste of expensive assets. Ramping up a signal in intensity may have no impact on geophysical operations, but sonar does not work very well if you give the target submarine a warning so that he can do countermeasures. On board civilian observers are used in tuna-boat operations, and in dredging exercises, which are radically different from military operations.

Chapter-18

Scientific Echosounder and Side-Scan Sonar

Scientific echosounder

A **scientific echosounder** is a device which uses SONAR technology for the measurement of underwater physical and biological components—this device is also known as *scientific SONAR*. Applications include bathymetry, substrate classification, studies of aquatic vegetation, fish, and plankton, and differentation of water masses.

Technology

Scientific echosounder equipment is built to exacting standards and tested to be stable and reliable in the transmission and receiving of sound energy under the water. Recent advances have led to the development of the digital scientific echosounder, further enhancing the reliability and precision with which these systems operate. Modern scientific echosounders are reliable, portable, and relatively easy to use.

Data collected with a scientific echosounder can be analyzed for the presence, abundance, distribution and acoustic characteristics of such variables as: depth (bathymetry), bottom substrate class (e.g., sand, mud, rock), submersed aquatic vegetation (SAV), and water column scattering (fish and plankton). Resulting analysis can be used to generate GIS data layers for these variables.

Applications

Scientific echosounders are commonly used by International, Federal, State and Local government and management agencies, as well as private-sector consultants working for these public agencies. Academic institutions have realized and are teaching the value of sampling non-invasively with sound to enhance both the spatial coverage and objectivity of fisheries sampling. Fisheries management agencies such as the membership of ICES and the United States National Marine Fisheries Service (NMFS) commonly use scientific sonar for stock assessment purposes, such as herring biomass assessment for resource management purposes.

More recently, the acoustic data collected has been valuable in underwater habitat assessment and classification for the variables; seabed type (e.g. rock, mud, sand) and submersed aquatic vegetation and algae - with the appropriate software.

Side-scan sonar

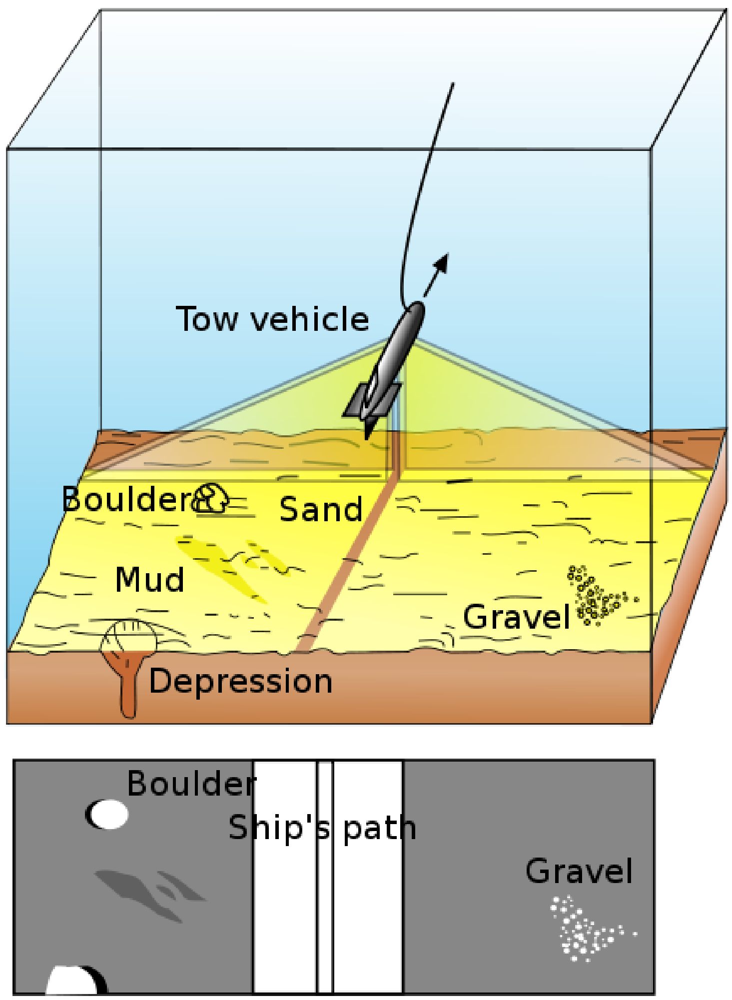
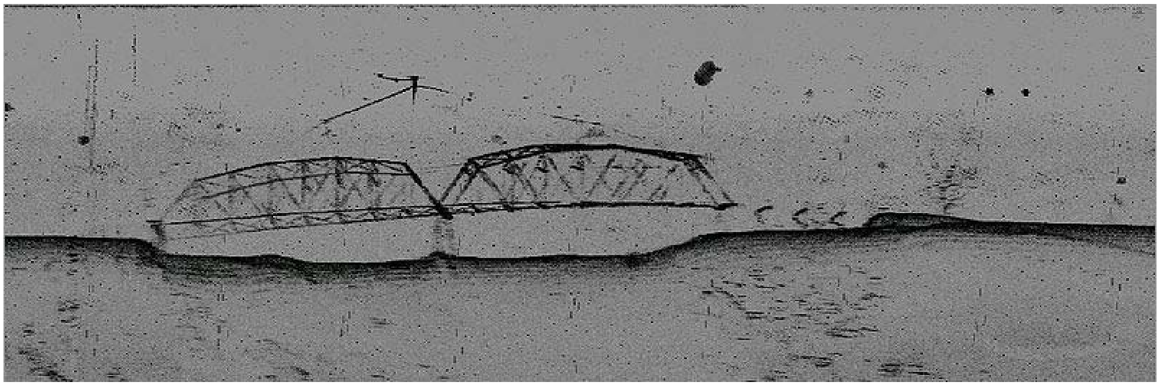


Diagram of sidescan sonar



Submerged bridge in 160 ft of fresh water seen on sidescan sonar imagery using a Humminbird 981c Side Imaging system

Side-scan sonar (also sometimes called **side scan sonar**, **sidescan sonar**, **side looking sonar**, **side-looking sonar**, **side imaging sonar**, **side-imaging sonar** and **bottom classification sonar**) is a category of sonar system that is used to efficiently create an image of large areas of the sea floor. It may be used to conduct surveys for maritime archaeology; in conjunction with seafloor samples it is able to provide an understanding of the differences in material and texture type of the seabed. Side-scan sonar imagery is also a commonly used tool to detect debris items and other obstructions on the seafloor that may be hazardous to shipping or to seafloor installations by the oil and gas industry. In addition, the status of pipelines and cables on the seafloor can be investigated using side-scan sonar. Side-scan data are frequently acquired along with bathymetric soundings and sub-bottom profiler data, thus providing a glimpse of the shallow structure of the seabed. Side-scan sonar is also used for fisheries research, dredging operations and environmental studies. It also has military applications including mine detection.

Side scan uses a sonar device that emits conical or fan-shaped pulses down toward the seafloor across a wide angle perpendicular to the path of the sensor through the water, which may be towed from a surface vessel or submarine, or mounted on the ship's hull. The intensity of the acoustic reflections from the seafloor of this fan-shaped beam is recorded in a series of cross-track slices. When stitched together along the direction of motion, these slices form an image of the sea bottom within the swath (coverage width) of the beam. The sound frequencies used in side-scan sonar usually range from 100 to 500 kHz; higher frequencies yield better resolution but less range.

The earliest side-scan sonars used a single conical-beam transducer. Next, units were made with two transducers to cover both sides. The transducers were either contained in one hull-mounted package or with two packages on either side of the vessel. Next the transducers evolved to fan-shaped beams to produce a better "sonogram" or sonar image. In order to get closer to the bottom in deep water the side-scan transducers were placed in a "towfish" and pulled by a "tow cable".

One of the inventors of side-scan sonar was German scientist, Dr. Julius Hagemann, who was brought to the US after World War II and worked at the US Navy Mine Defense

Laboratory, Panama City, FL from 1947 until his death in 1964. His work is documented in US Patent 4,197,591 which was first disclosed in Aug 1958, but remained classified by the US Navy until it was finally issued in 1980. Experimental side-scan sonar systems were made during the 1950s in laboratories including Scripps Institution of Oceanography and Hudson Laboratories and by Dr. Harold Edgerton at MIT.

Military side-scan sonars were made in the 1950s by Westinghouse. Advanced systems were later developed and built for special military purposes, such as to find H-Bombs lost at sea or to find a lost Russian submarine, at the Westinghouse facility in Annapolis up through the 1990s. This group also produced the first and only working *Angle Look Sonar* that could trace objects while looking under the vehicle.

The first commercial side-scan system was the Kelvin-Hughes "Transit Sonar", a converted echo-sounder with a single-channel, pole-mounted, fan-beam transducer introduced around 1960. In 1963 Dr. Harold Edgerton, Edward Curley, and John Yules used a conical-beam 12kHz side-scan sonar to find the sunken Vineyard Lightship in Buzzards Bay, Massachusetts. A team led by Martin Klein at Edgerton, Germeshausen & Grier (later E.G. & G., Inc.) developed the first successful towed, dual-channel commercial side-scan sonar system from 1963 to 1966. Martin Klein is generally considered to be the "father" of commercial side-scan sonar. In 1967, Edgerton used Klein's sonar to help Alexander McKee find Henry VIII's flagship *Mary Rose*. That same year Klein used the sonar to help archaeologist George Bass find a 2000-year-old ship off the coast of Turkey. In 1968 Klein founded Klein Associates, Inc.(now L-3/Klein) and continued to work on improvements including the first commercial high frequency (500 kHz) systems and the first dual-frequency side-scan sonars, and the first combined side-scan and sub-bottom profiling sonar. In 1985, Charles Mazel of Klein Associates produced the first commercial side-scan sonar training videos and the first *Side Scan Sonar Training Manual*.

For surveying large areas, the GLORIA sidescan sonar was developed by Marconi Underwater Systems for NERC. This operated at relatively low frequencies to obtain long range. It was used by the US Geological Survey and the Institute of Oceanographics in the UK to obtain images of continental shelves world-wide.

Manufacturers of higher frequency side-scan sonar systems include Lowrance, Simrad, Raytheon, Northrop Grumman (formerly Westinghouse), EdgeTech (formerly E.G. & G.), C-MAX Ltd., L-3/Klein Associates, J.W. Fishers Mfg. Inc., Imagenex Technology Corp., RESON A/S, Sonatech Inc., Benthos (the sonar formerly produced by Datasonics), WESMAR, Marine Sonic Technology, Kongsberg Maritime, Geoacoustics, EDO Corp., Ultra Electronics,Humminbird (Techsonic Industries Inc) and Deep Vision Technologies.

Up until the mid 1980s, commercial sidescan images were produced on paper records. The early paper records were produced with a sweeping plotter that burned the image into a scrolling paper record. Later plotters allowed for the simultaneous plotting of position and ship motion information onto the paper record. In the late 1980s, commercial systems using the newer, cheaper computer systems developed digital scan-converters that could mimic more cheaply the analog scan converters used by the military systems to produce TV and computer displayed images of the scan, and store them on video tape. Now data is stored on computer hard drives.

Chapter-19

Seismic Source

An air gun seismic source (30 litre)

A **seismic source** is a device that generates controlled seismic energy used to perform both reflection and refraction seismic surveys. A seismic source can be simple, such as dynamite, or it can use more sophisticated technology, such as a specialized air gun. Seismic sources can provide single pulses or continuous sweeps of energy that generate seismic waves, which travel through a medium such as water or layers of rocks. Some of the waves then reflect and refract and is recorded by receivers, such as geophones or hydrophones.

Seismic sources may be used to investigate shallow subsoil structure, for engineering site work, or deeper structures, usually in the search for petroleum or mineral deposits, or for scientific investigation. The returning signals from the sources are detected by geophones, laid in known locations relative to the position of the source. The recorded signals are then subjected to specialist processing and interpretation to yield comprehensible data about the subsurface.

Source model

A seismic source signal has the following characteristics:

1. generated as an impulsive source
2. band-limited
3. the generated waves are time-varying

The generalized equation that shows all above properties is:

$$s(t) = \beta e^{-\alpha t^2} \sin(2\pi f_{max} t)$$

where f_{max} is the maximum frequency component of the generated waveform.

Types of sources

Explosives

Explosives, such as dynamite, can be used as crude but effective sources of seismic energy. Generally the explosive charges are placed between 20 feet to 250 feet below ground. The charges are placed in a hole that is drilled with dedicated drilling equipment for this purpose. This type of seismic drilling is often referred to as "Shot Hole Drilling".

A common drill rig used for "Shot Hole Drilling" is the ARDCO C-1000 drill mounted on an ARDCO K 4X4 buggy. These drill rigs often use water or air in assisting the drilling.

Air gun

An **air gun** is used for marine reflection and refraction surveys. It consists of one or more pneumatic chambers that are pressurized with compressed air at pressures from 2,000 pounds per square inch to 3,000 pounds per square inch (14 to 21 MPa). The air gun array is submerged below the water surface, and is towed behind a ship. When the air gun is fired, a solenoid is triggered, which releases air into a fire chamber which in turn causes a piston to move and thereby allowing the air to escape the main chamber and to produce a pulse of acoustic energy. Air gun arrays are built up of up to 48 individual air guns with different size chambers, the aim being to create the optimum initial shock wave with minimum reverberation of the bubble after the first shot.

Gun arrays can be fired in flip-flop mode; typically this would be 48 guns per source, which would be selected and fired alternately. Large chambers (i.e., greater than 70 cubic inches or 1.15 L) tend to give low frequency signals, and the small chambers (less than 70 cubic inches) give higher frequency signals. The air gun is made from the highest grades of corrosion resistant stainless steel.

Plasma sound source

Plasma sound source fired in small swimming pool

A **plasma sound source** (PSS), otherwise called a **spark gap sound source**, or simply a **sparker**, is a means of making very low frequency sonar pulse underwater.

For each firing, it stores electric charge in a large high-voltage bank of capacitors, and then releases all the stored energy in an arc across electrodes in the water. The underwater spark discharge produces a high-pressure plasma and vapor bubble, which expands and collapses, making a loud sound. Most of the sound produced is between 20 and 200 Hz.

The PSS has also been used for sonar. There are also plans to use PSS as a non-lethal weapon against submerged divers.

Thumper truck

A **thumper truck** (or weight-drop) truck is a vehicle mounted ground impact which can used to provide the seismic source. A heavy weight is raised by a hoist at the back of the truck and dropped, possibly about three metres, to impact (or "thump") the ground. To augment the signal, the weight may be dropped more than once at the same spot, the signal may also be increased by thumping at several nearby places in an array whose dimensions may be chosen to enhance the seismic signal by spatial filtering.

Thumping might be less damaging to the environment than firing explosives in shot-holes, though a heavily thumped seismic line with transverse ridges every few metres might create long-lasting disturbance of the soil. An advantage of the thumper (later shared with Vibroseis), especially in politically unstable areas, was that no explosives were required.

Seismic vibrator

A Seismic vibrator, commonly known by its trademark name **Vibroseis**, propagates energy signals into the Earth over an extended period of time as opposed to the near instantaneous energy provided by impulsive sources. The data recorded in this way must be *correlated* to convert the extended source signal into an impulse. The source signal using this method was originally generated by a servo-controlled hydraulic vibrator or *shaker unit* mounted on a mobile base unit, but electro-mechanical versions have also been developed.

Vibroseis was developed by the Continental Oil Company (Conoco) during the 1950s and was a trademark until the company's patent lapsed.

Boomer sources

Boomer sound sources are used for shallow water seismic surveys, mostly for engineering survey applications. Boomers are towed in a floating sled behind a survey vessel. Similarly to the plasma source, it stores energy in capacitors, but it discharges through a flat spiral coil instead of generating a spark. A copper plate adjacent to the coil flexes away from the coil as the capacitors are discharged. This flexing is transmitted into the water as the seismic pulse.

Originally the storage capacitors were placed in a steel container (the **bang box**) on the survey vessel. The high voltages used, typically 3,000 V, required heavy cables and strong safety containers. Recently, low voltage boomers have become available. These use capacitors on the towed sled, allowing efficient energy recovery, lower voltage power supplies and lighter cables. The low voltage systems are generally easier to deploy and have fewer safety concerns.

Noise sources

Correlation-based processing techniques also enable seismologists to image the interior of the Earth at multiple scales using natural (e.g., the oceanic microseism) or artificial (e.g., urban) background noise as a seismic source. For example, under ideal conditions of uniform seismic illumination, the correlation of the noise signals between two seismographs provides an estimate of the bidirectional seismic impulse response.